Allosteric Regulation of the Adenosine Triphosphate Phosphoribosyltransferase from *Campylobacter jejuni*

A thesis submitted in partial fulfilment
of the requirements for the degree
of
Doctor of Philosophy in Biochemistry
in the Department of Chemistry
by Gerd Horst Mittelstädt

May 2015

Abstract

The enzyme adenosine triphosphate phosphoribosyltransferase (ATP-PRT) catalyses the first reaction of the histidine biosynthetic pathway. ATP-PRT also represents a metabolic control point, directing the flux of metabolites through this energetically expensive pathway. Two distinctly different forms of ATP-PRT exist, the long form and the short form, which differ in the presence of a C-terminal regulatory domain. In the short form, where this domain is absent, it is functionally replaced by a regulatory protein, called HisZ. ATP-PRT activity is modulated by two layers of regulation: active site inhibition by adenosine monophosphate, which reflects cellular energy levels, and pathway end product feedback inhibition by histidine. In the long form ATP-PRT histidine binds to the allosteric site at the regulatory domain, but the exact nature of the inhibitory mechanism is still debated.

This thesis characterises a new member of the ATP-PRT long form from *Campylobacter jejuni* (*Cje*ATP-PRT) and investigates the molecular mechanisms involved in the feed back inhibition by histidine.

Chapter 2 describes the characterisation of the *Cje*ATP-PRT including a detailed description of its crystal structure. The *C. jejuni* enzyme is similar to the previously described enzymes of the ATP-PRT long form, but exists only as hexameric species under experimental conditions, which contradicts previous assumptions that the hexamer is exclusively inactive.

Chapter 3 investigates the catalytic apparatus of *Cje*ATP-PRT by separating the catalytic and regulatory domains of the enzyme for individual study. The isolated catalytic portion of the enzyme, the *Cje*ATP-PRT Core

mutant, forms a dimeric species with very limited catalytic capabilities but high substrate and product affinities. The *Cje*ATP-PRT Core characteristics suggest that it exists in a permanently inhibited conformation, highlighting the requirement of the regulatory domain not only for feedback regulation but also for enzyme function. Additionally this supports the evolutionary need for the recruitment of a regulatory apparatus.

In chapter 4 a potential intramolecular communication pathway from the allosteric to the active site is probed by the generation of several single site mutations. One of these, *Cje*ATP-PRT R216A, is completely insensitive to histidine inhibition, although this ligand is still able to bind at the allosteric site, which is consistent with the involvement of R216 in the allosteric signal communication. The catalytic abilities of *Cje*ATP-PRT R216A are largely impaired, leading to the assumption that this mutation causes a permanent inhibitory response.

In summary this thesis supports the existence of a simple physical regulatory mechanism for the feedback inhibition of the ATP-PRT long form, the change between two different hexamer conformations depending on the presence of the allosteric effector.

Acknowledgments

I want to give thanks to everyone who directly and indirectly influenced the existence of this thesis. First of all to my supervisor, Emily Parker. I can not emphasise enough how grateful I am for her support and encouragement over the years. She took me on as a research assistant, when I was plan- and guideless, and groomed me to a stage where I could see myself as a successful PhD candidate. Her enthusiasm for her profession is both infectious and inspiring and with her accumulated knowledge and experience she always spots a way around, or if necessary through, any obstacles that might come along. And while she works ridiculous hours to achieve everything, my deepest gratitude is with her. Thanks for everything, Emily.

I thank the University of Canterbury for providing me with a UC Doctoral Scholarship to undertake my research and the Biomolecular Interaction Centre for technical, intellectual and additional financial support throughout this thesis.

Big thanks go to Gert-Jan and Alicia for letting me stay with them in the dire times when the rest of my little family had already departed New Zealand. Closely associated with my PhD project, Gert-Jan also had a large influence on the crystallographic successes presented in this thesis.

My thanks also goes to Tim for teaching me all the tricks in and around the lab when I first started. Tim taught me the basics of crystallography and always new the magic tweak for tricky refinements. He also had a large input during the write up of this thesis, providing almost a 'life update' from the other side of the world.

I want to give a special thanks to Penel for her strong support during my PhD years. Her outright honesty and keen eye for the important things always kept me on the right track and her hidden stashes often saved the day. Thanks to the hospitality of Penel and her family it is possible for me to enjoy my remaining time in New Zealand without worries.

Amidst all the chaos of trying to finish up, Nicky has been of great help to me. Whenever I had a problem, be it data processing, bent ITC needles, broken chromatography systems or retrieving seven mm long magnetic fleas from a full 240 L rubbish bin, I could always count on her.

To my "Lab-hand" Haley I owe the existence of chapter 4. Her sharp mind and steady hands were an amazing and invaluable support during the last months of my PhD lab work.

Sincere thanks go to Emma for voluntarily reading my entire thesis and suggesting valuable corrections.

I also thank all other current and past members of the Parker group, especially Richard, Sebastian, Sarah, and Logan for their friendship, always offering me a helping hand, and the productive (and less productive) discussions in and outside the lab, over the past years. It has been a great journey with you guys.

Finally, thanks to my family, especially my wife Jana, for their incredible support and the more than valuable moments of retreat during the years of my PhD. Jana has managed the life of our little but steadily growing family in New Zealand. From the day-to-day business to the big events and decisions, she has it all planned out well in advance. On the way, we also raised two amazing little boys. You truly define who I am and I could not have done this without you, thank you.

Abbreviations

ADP	adenosine diphosphate
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AMP	adenosine monophosphate
*Ath*ATP-PRT	*Arabidopsis thaliana* ATP-PRT
ATP	adenosine triphosphate
ATP-PRT	adenosine triphosphate phosphoribosyltransferase
AUC	analytical ultra centrifugation
BTP	1,3-bis[tris(hydroxymethyl)methylamino]propane
CD	circular dichroism
*Cgl*ATP-PRT	*Corynebacterium glutamicum* ATP-PRT
*Cje*ATP-PRT	*Campylobacter jejuni* ATP-PRT
DNA	deoxyribonucleic acid
DSC	differential scanning calorimetry
DSF	differential scanning fluorimetry
DTT	dithiothreitol
*Eco*ATP-PRT	*Escherichia coli* ATP-PRT
*Eco*PPase	*E. coli* PPase
EDTA	ethylenediaminetetraacetic acid
GST	glutathione sulfur-transferase
Hol-P	histidinol phosphate
IAP	imidazole acetolphosphate
IGP	imidazole glycerol phosphate
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside

ITC	isothermal titration calorimetry
LB	lysogeny broth
*Lla*ATP-PRT	*Lactococcus lactis* ATP-PRT
*Mth*ATP-PRT	*Methanothermobacter thermautotrophicus* ATP-PRT
*Mtu*ATP-PRT	*Mycobacterium tuberculosis* ATP-PRT
MWCO	molecular weight cut off
NAD	nicotinamide adenine dinucleotide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PP	pyrophosphate
PPase	pyrophosphatase
PR-AMP	phosphoribosyl-adenosine monophosphate
PR-ATP	phosphoribosyl-adenosine triphosphate
PRFAR	phosphoribulosyl-formimino-AICAR-phosphate
ProFAR	phosphoribosyl-formimino-AICAR-phosphate
PRPP	phosphoribosyl pyrophosphate
PRT	phosphoribosyltransferase
R5P	ribose 5-phosphate
RMSD	root-mean-square deviation
RNA	ribonucleic acid
SAXS	small angle X-ray scattering
SDM	site-directed mutagenesis
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
*Sen*ATP-PRT	*Salmonella enterica* subsp. *enterica* Typhimurium ATP-PRT
SLS	static light scattering
SOC	super optimal broth
TAE	Tris-acetate-EDTA
TEV	tobacco etch virus
*Tma*ATP-PRT	*Thermotoga maritima* ATP-PRT
Tris	tris(hydroxymethyl)aminomethane

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Histidine and the histidine biosynthetic pathway

L-Histidine is one of the 22 proteogenic amino acids, the basic metabolites from which proteins are synthesised in the fundamental process of translation. The functional group of histidine, imidazole, is able to be protonated and deprotonated at physiological pH, which makes it an excellent proton shuttle during enzymatic reactions. Histidine is therefore often found in enzyme active sites, for instance as part of the catalytic triad of peptidases.[1] The side chain of histidine is also a very common coordinating ligand in metalloproteins.[2,3] Aside from the roles of histidine in protein structure and function, it can be decarboxylated to form the common neurotransmitter histamine.

Histidine is one of the essential amino acids that humans require as part of their diet. Its biosynthesis is ubiquitous, occurring in all domains of life except animals, which makes the enzymes of the histidine biosynthetic pathway interesting targets for the development of non-toxic drugs and herbicides.[4,5]

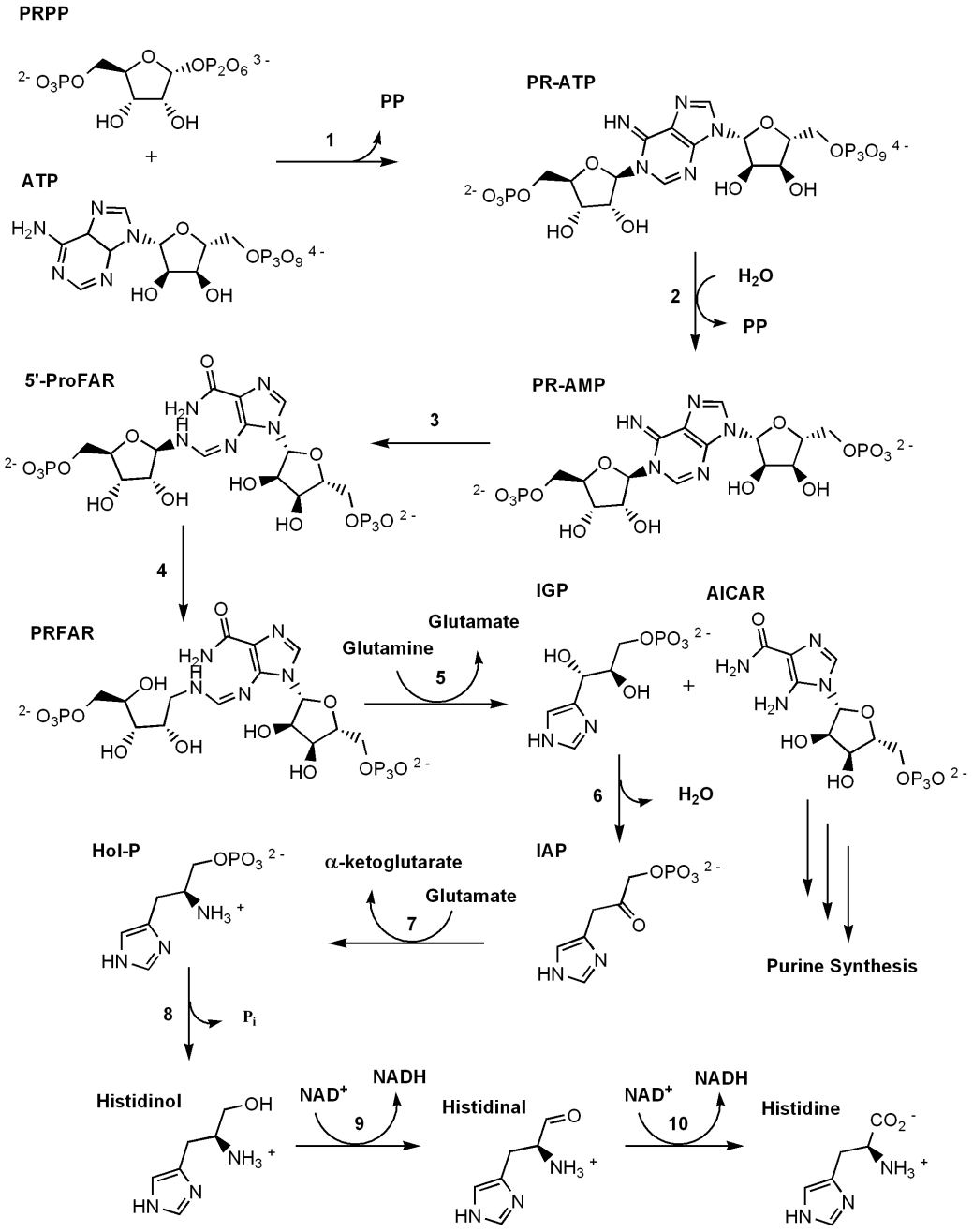


Figure 1.1: Biosynthetic pathway of the proteogenic amino acid histidine. Overview of the ten reaction steps involved in the biosynthetic production of histidine. Participating enzymes: 1 – ATP-PRT, 2 – PR-ATP pyrophosphohydrolase, 3 - PR-AMP cyclohydrolase, 4 – ProFAR isomerase, 5 – IGP synthase/glutaminase, 6 – IGP dehydratase, 7 – IAP aminotransferase, 8 - Hol-P phosphatase, 9 and 10 – histidinol dehydrogenase.

The investigation of the histidine biosynthetic pathway began in 1952 with studies on *Neurospora crassa* mutants[6] and soon included bacteria[7,8] and later plants.[9] The work on the histidine biosynthetic pathway has largely contributed to the unravelling of some of the most fundamental mechanisms in biology during the 20th century, such as operons[10] and operon polarity,[11] attenuation[12] and feedback inhibition.[13] Today a clear picture of the reactions that lead to the formation of histidine, the involved metabolites and enzymes, in all domains of life, has been obtained.[5,9,14]

Histidine biosynthesis is achieved in ten enzymatic reaction steps (Figure 1.1). Using adenosine triphosphate (ATP) and phosphoribosyl pyrophosphate (PRPP), from the pentose phosphate pathway, first phosphoribosyl-adenosine triphosphate (PR-ATP) is formed by a phosphoribosyl transfer reaction (step 1) catalysed by adenosine triphosphate phosphoribosyltransferase (ATP-PRT). The subsequent hydrolysis of the triphosphate (step 2) is catalysed by PR-ATP pyrophosphohydrolase, which results in phosphoribosyl-adenosine monophosphate (PR-AMP). Hydrolytic opening of the adenosine ring (step 3) produces phosphoribosyl-formimino-AICAR-phosphate (ProFAR). This reaction is catalysed by PR-AMP cyclohydrolase. Depending on the organism, PR-ATP pyrophosphohydrolase and PR-AMP cyclohydrolase occur as two separate monofunctional enzymes, a fused bifunctional enzyme or a single multifunctional enzyme.[5] ProFAR undergoes ribose ring opening by way of an Amadori rearrangement (step 4) to produce phosphoribulosyl-formimino-AICAR-phosphate (PRFAR), which is then converted into imidazole glycerol phosphate (IGP) and 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) (step 5) by IGP synthetase/glutaminase. This enzyme is either a non-covalent complex[14] (prokaryotes) or a fused bifunctional gene product[15] (eukaryotes) and provides the ammonia required for the indole ring formation via the deamination of glutamine. The liberation of AICAR also connects the histidine biosynthetic pathway to the *de novo* purine biosynthesis, as this compound is a key entry point in the purine pathway. IGP is further dehydrated (step 6) to yield imidazole acetolphosphate (IAP) by IGP dehydratase. After transamination (step 7) of IAP to histidinol phos-

phate (Hol-P) by IAP aminotransferase, the Hol-P phosphatase catalyses the dephosphorylation (step 8) of Hol-P to histidinol. In eubacteria IGP dehydratase and Hol-P phosphatase can also appear as a single bifunctional enzyme. Histidinol is finally oxidised by histidinol dehydrogenase in two sequential steps via histidinal (step 9 and 10) to form the pathway end-product histidine.

1.2 The adenosine triphosphate phosphoribosyltransferase

ATP-PRT (EC 2.4.2.17) catalyses the first reaction of histidine biosynthesis, the reversible transfer of the 5'-phosphoribosyl group of PRPP on to the N_1 atom of ATP, resulting in the products PR-ATP and PP (Figure 1.2). The reaction is described as an equilibrium and Mg^{2+} is required for optimal enzymatic efficiency.[16] ATP-PRT is strongly inhibited by the pathway end product histidine in both directions[13,17] and also competitively inhibited by the nucleotides adenosine monophosphate (AMP) and to a lesser extent adenosine diphosphate (ADP). Combinations of AMP or ADP with histidine have been shown to exhibit synergistic inhibition towards ATP-PRT.[18]

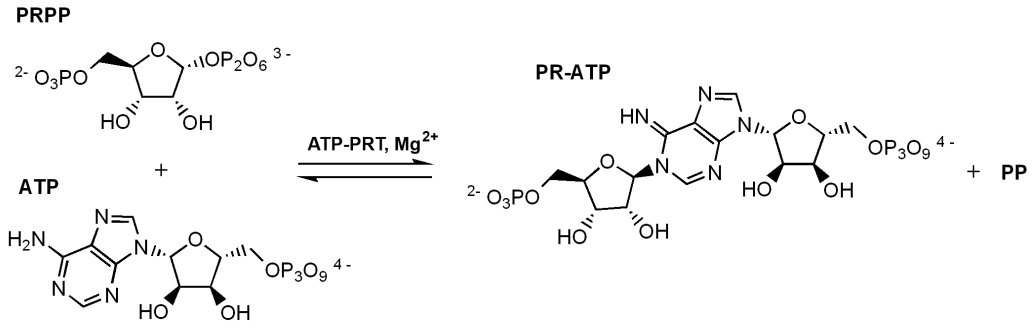


Figure 1.2: The ATP-PRT reaction.

As mentioned, histidine biosynthesis, as a central metabolic pathway, is connected to the purine *de novo* synthesis by the pathway side product AICAR, which is released upon the formation of IGP. Accumulation

of AICAR has also been shown to adversely affect the synthesis of thiamine pyrophosphate in *Salmonella enterica*[19] and tetrahydrofolate in *Saccharomyces cerevisiae*,[20] two important metabolites. Furthermore, the substrates of ATP-PRT are both energetically rich compounds, and play important roles in cellular metabolism. It has been estimated that the synthesis of one molecule of histidine requires a total of 41 equivalents of ATP.[21] It is therefore essential that histidine biosynthesis is tightly regulated. Catalysing the first committed step of the pathway, ATP-PRT plays a key role in the regulation of metabolite flux through the pathway. It is consequently controlled by multiple layers of regulation, which are described in detail later (section 1.3).

1.2.1 Classification of phosphoribosyltransferases

ATP-PRT is a member of the phosphoribosyltransferase (PRT) enzyme family. Four types of PRTs are differentiated according to their overall fold (Figure 1.3). The most common of these is type I, which includes enzymes involved in nucleotide synthesis and salvage pathways. The typical type I fold consists of a central five-stranded β sheet surrounded by three α helices, accompanied by a predominantly N-terminal, variable "hood" domain. In most examples of type I PRTs two individual chains form a dimeric interface that includes residues of the active site.[22] Quinolate PRT represents another type, type II, and its chain is composed of two domains, an N-terminal antiparallel β openface sandwich domain and an uncommon $\beta 7 \alpha 6$ barrel.[23] Two chains of quinolate PRT come together to form the holo-enzyme. The only representative of type III, anthranilate PRT, also adopts a stable dimeric quaternary structure. The type III single chain is composed of a small N-terminal four α helix bundle and a large C terminal α/β domain with a central seven stranded β sheet.[24,25] Notably, ATP-PRT is structurally dissimilar to all other PRT family members, making it a separate type: type IV PRT.[26] ATP-PRT enzymes can be further subdivided into two forms, a homo-hexameric and a hetero-octameric form. The common feature of

both forms are two α/β sandwich domains that form the active site in the large cleft between them. The detailed structure of ATP-PRT is described in section 1.2.3.

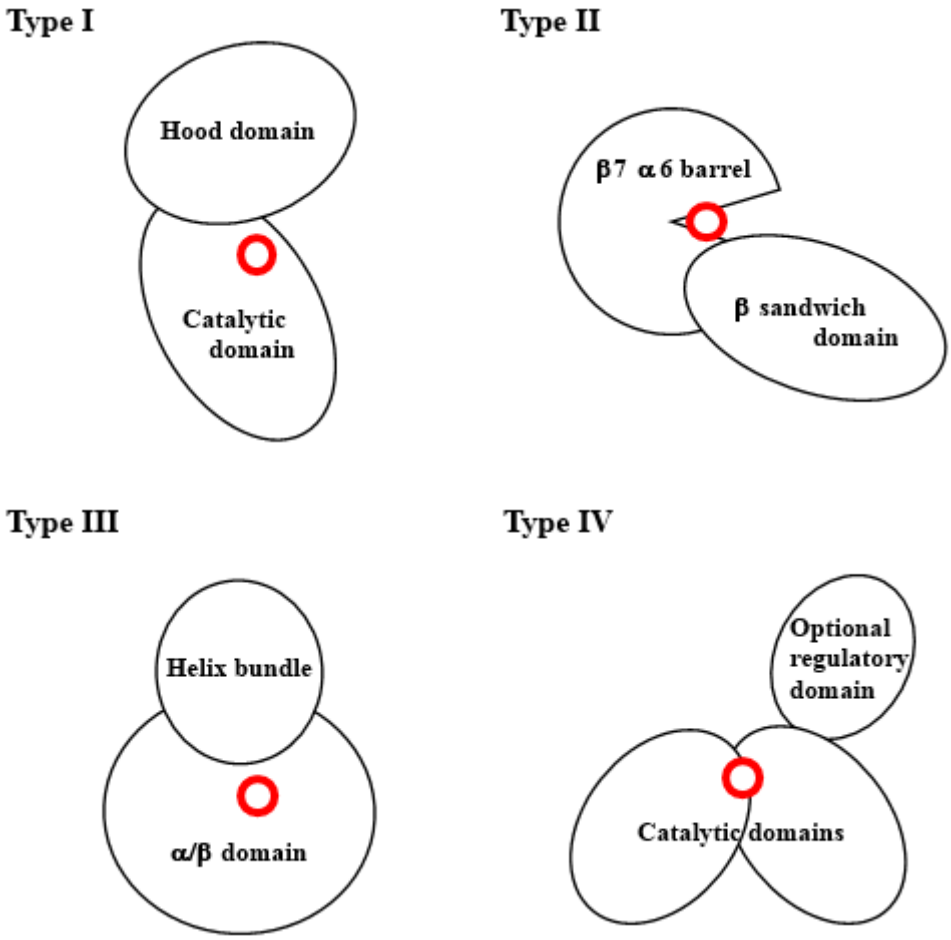


Figure 1.3: Overview of the PRT type architectures. Schematic comparison of PRT type structures.[22–24,35] For each type the domains of a single chain are depicted as ovals with the position of the active site indicated by a red circle. Type IV represents the ATP-PRT family.

A 13 residue signature fold responsible for PRPP binding has been described for the type I PRTs. The consensus sequence of this highly conserved PRPP binding motif can be generally described as hhhhaahhssss, where 'h' represents a hydrophobic side chain, 'a' an acidic side chain (usually aspartic acid) and 's' a short side chain. The first four hydrophobic residues

form the central β strand of the five-stranded β sheet, while the last nine residues of the motif compose the PRPP binding loop. The two acidic side chains lie in hydrogen bonding distance from the ribose oxygens of PRPP and the 5'-phosphate group binds in a positively charged pocket formed by backbone amides and polar side chains of the last five residues, including at least one glycine and one threonine residue. This binding loop is very potent and specific for the binding of negatively charged moieties or molecules. In crystal structures of PRT proteins, crystallised in the absence of specific ligands, the binding site formed by this loop is nearly always occupied by anions present in the crystallisation condition, such as sulphate, phosphate, or sulphonate.[22] Type II and III PRTs do not contain sequences similar to this conserved motif, but have instead developed separate coordination sites for the binding of PRPP.[23,25] In type IV, which is all ATP-PRTs, the same signature motif is found but with minor differences. Here, the second residue in the sequence is acidic while position five is always a hydrophobic residue, resulting in a consensus sequence, using the same nomenclature, of hahh-hahhssss. The last small residue is almost exclusively a threonine.[27] Similar to the type I, the first five residues form a β sheet (β9), while the remaining residues form the binding loop.[26] When the two 13 amino acid motifs from the crystal structures of type I *Trypanosoma cruzi* hypoxanthine PRT and type IV *Escherichia coli* ATP-PRT (*Eco*ATP-PRT) are superimposed, they display a striking similarity (Figure 1.4), with a root-mean-square deviation (RMSD) of 0.895 Å, showing that the PRT binding motif is almost identical in sequence and fold in the two PRT types.

1.2.2 Reaction mechanism

Extensive kinetic studies on the ATP-PRT from *Salmonella enterica* subsp. *enterica* Typhimurium (*Sen*ATP-PRT) have established that the ATP-PRT reaction proceeds via an ordered sequential binding mechanism with ATP binding first, followed by PRPP, release of PP, and release of PR-ATP.[17,28]

The reaction catalysed by ATP-PRT is a nucleophilic substitution with

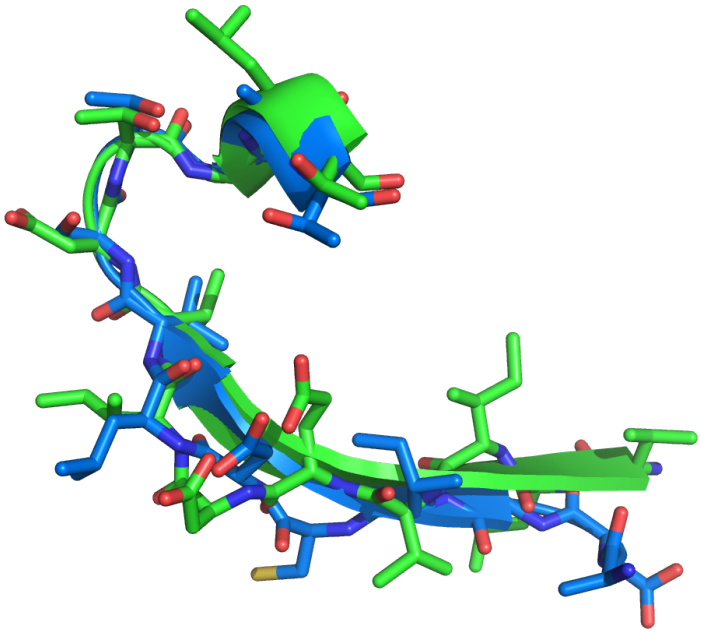


Figure 1.4: The PRPP binding motif. Superposition of the PRPP binding motif of *T. cruzi* hypoxanthine PRT (green) and *Eco*ATP-PRT (marine) in the cartoon and stick representations. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), sulphur (yellow).

inversion of stereochemistry.[29] Although the mechanism of the ATP-PRT reaction has not yet been determined, two mechanisms can be proposed based on the reaction chemistry. The first is a dissociative (S_N1-like) mechanism, in which the pyrophosphate of PRPP dissociates first, creating an oxocarbenium ion intermediate, which is then attacked by the nucleoplilic N_1 atom of ATP. The second is an associative (S_N2-like) mechanism, in which nucleophilic attack of ATP and dissociation of the pyrophosphate group occur simultaneously, without the formation of a discrete intermediate.

The determination of the transition state structure can discriminate the degree of S_N1-like or S_N2-like character of a reaction.[30] Studies of the kinetic isotope effect on a variety of PRT enzymes including *Sen*ATP-PRT

have shown that most of the investigated PRT reactions follow an S_N1-like mechanism. The observed large ^3H kinetic isotope effects are consistent with the formation of an oxocarbenium ion intermediate. However the results for *Sen*ATP-PRT were less conclusive, which could be due to the bi-directional character of the reaction.[31]

Analogies can also be made to the mechanisms of PRT proteins in general. The majority of the information available is concerning type I proteins, where crystallographic data shows that the ribose ring of PRPP is bound in the active site in a conformation that promotes a planar geometry at C_1. This is consistent with a partial double bond character in the C_1-O_4 bond, as would be expected with the formation of an oxocarbenium ion intermediate. Additionally an α-phosphate oxygen of the pyrophosphate group is in close proximity to C_1, which would further stabilise the planar intermediate.[32] Taken together this supports a dissociative mechanism of the type I PRT. Furthermore, it has been shown that the ribose conformation of PRPP is variable depending on whether it is bound to PP (substrate side) or ATP (product side), which is thought to be consistent with a general S_N1-like mechanism for the reaction of PRT enzymes.[33]

Recent investigation into the anthranilate PRT from *Mycobacterium tuberculosis* (type III PRT) revealed the role of flexible active site loops in the binding of substrate molecules, where the loops can move to effectively occlude the active site. This is thought to help protect the hypothesised highly reactive oxocarbenium ion intermediate. Additionally, kinetic studies of this enzyme with fluorinated substrate analogues indicate that the nucleophilicity of the attacking amino group (of anthranilate) is of little significance to the overall reaction rate, further supporting an S_N1-like mechanism for PRTs.[34] It is therefore generally assumed that the ATP-PRT reaction also displays an S_N1-like mechanism, but detailed analysis of the transition state involved are still under way (Gerard-Johan Moggré, personal communication, April 2015).

In crystal structures of almost all substrate bound PRT enzymes, Mg^{2+} ions have been modelled, positioned such that they would be involved in

PRPP binding and recognition.[22,25] It has also been shown that the presence of Mg^{2+} in *in vitro* assays is essential for ATP-PRT activity.[16] While it is certain that Mg^{2+} forms a complex with the polyphosphate groups of both ATP-PRT substrates,[28] a clear role of the metal ion during the catalytic mechanism has not been described.

1.2.3 Structure

Crystal structures have been solved of ATP-PRT from a variety of bacterial sources, including *M. tuberculosis*,[35] *E. coli*,[26] *Lactococcus lactis*[36] and *Thermotoga maritima*.[37] The Protein Data Bank (PDB) also contains the ATP-PRT crystal structure from the archaea *Methanothermobacter thermautotrophicus* (*Mth*ATP-PRT). Interestingly, from analysis of these structures it is clear that there are two distinct forms of ATP-PRT differentiated by their quaternary structure. Discrimination between the two forms starts at the primary sequence level. The first form is 280-310 amino acids in length, and is present in archaea, bacteria and eukaryotes. It will henceforth be denoted as "long form" ATP-PRT. The polypeptide of the second form, the "short form" ATP-PRT, consists of only 200–220 amino acid residues. The short form ATP-PRT has only been identified in bacteria and associates with a second gene product, HisZ, to form a functional complex. Of the above mentioned organisms, only *T. maritima* and *L. lactis* possess short form enzymes, denoted *Tma*ATP-PRT and *Lla*ATP-PRT respectively.

The polypeptide chain of ATP-PRT long form enzymes[26,35] is approximately 280–310 amino acids long containing a total of ten α helices and 15 β strands that form three discrete domains (Figure 1.5), of which domain I and II are involved in substrate binding and catalysis. The entity of both domains will be referred to as the catalytic core. Domain I is formed from two residue ranges, which are discontinuous, interrupted by the insertion of domain II between them. It consists of a six-stranded anti-parallel β sheet ($\beta 2 \beta 3 \beta 1 \beta 4 \beta 11 \beta 5$) flanked by four α helices, $\alpha 2$ and $\alpha 3$ on one side and $\alpha 1$ and $\alpha 8$ on the other, with the long $\alpha 8$ helix forming the connection to the

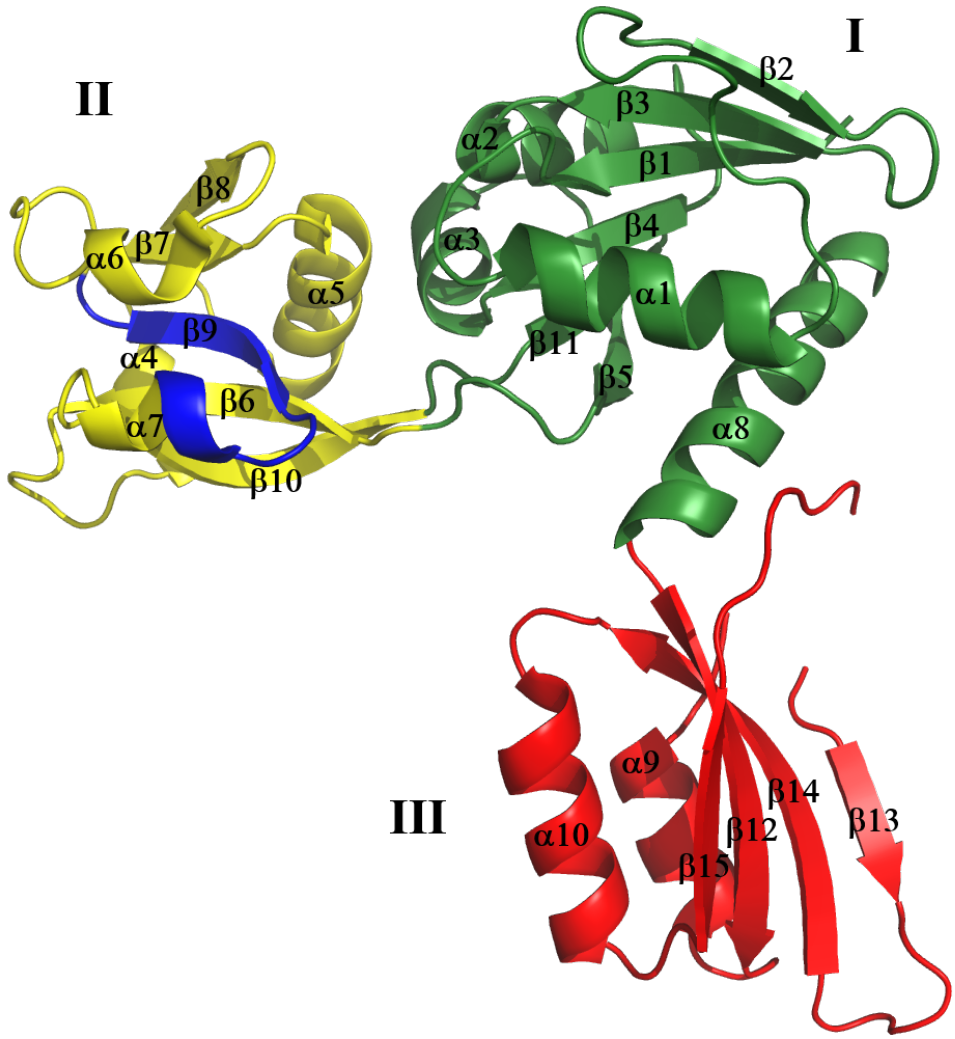


Figure 1.5: ATP-PRT long form structure. The crystal structure of apo *Mtu*ATP-PRT (PDB code: 1NH7) displayed in the cartoon representation. The contiguous domains I (green) and II (yellow) are connected via two loops, the C-terminal domain III (red) is linked to domain I via a single loop. α helices and β strands are labelled in primary sequence order. The PRPP binding motif is highlighted in blue.

C-terminal domain III. The fold of domain II is similar to domain I, and is composed of a twisted central β sheet ($\beta 8 \beta 7 \beta 9 \beta 6 \beta 10$) surrounded by two α helices on either side, $\alpha 4$ plus $\alpha 5$ and $\alpha 6$ plus $\alpha 7$ respectively. The cleft between domain I and II forms the active site, which includes the characteristic PRPP binding motif.

A very similar fold of domain I and II can also be found for the members of the ATP-PRT short form,[36,37] but with $\alpha 8$ being shorter (Figure 1.6).

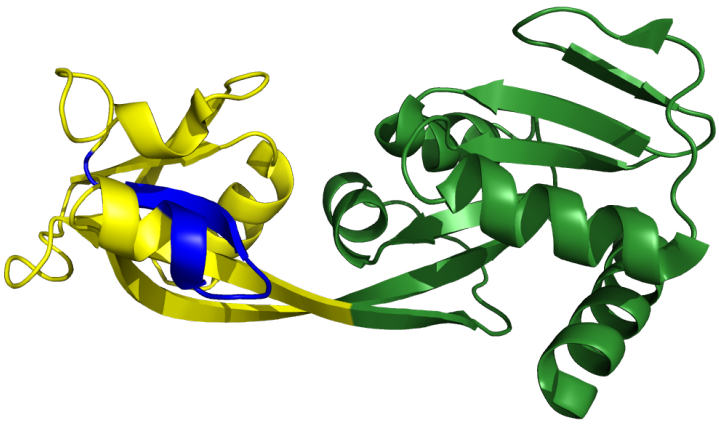


Figure 1.6: ATP-PRT short form structure. The crystal structure of apo *Tma*ATP-PRT (PDB code: 1USY) displayed in the cartoon representation and coloured according to Figure 1.5.

The C-terminal domain III, absent in the ATP-PRT short form, is also a mixed α/β domain. The five β strands and two α helices form an anti-parallel β sheet with both α helices on one face (Figure 1.5). This typical $\beta\alpha\beta\beta\alpha\beta$, ferredoxin-like fold has been described as the structural characteristic of the ACT-domain,[38] named after three enzymes, in which it was first identified: aspartate kinase, chorismate mutase and TyrA (prephenate dehydrogenase).[39] ACT domains are structural elements commonly found in allosterically regulated metabolic enzymes, predominantly involved in amino acid synthesis, but the ACT domain has also been found in DNA-binding proteins.[40] Because of the relative abundance of ferredoxin-like folds, ACT domains have been further defined by a combination of structure and function. Characteristic features of ACT domains include highly varied modes of self-association, effector binding at ACT domain interfaces, and transmission of allosteric signals.[41] Due to their involvement in allostery, ACT domains are considered "regulatory switches" of enzymes. The ACT domain of long form ATP-PRT enzymes is responsible for the binding of histidine and provides a key interface for the oligomeric assembly of the protein.

1.2.4 Macromolecular assembly

The quaternary structure also differs drastically between the ATP-PRT long form and short form enzymes. Long form ATP-PRT enzymes are described to exist in two homo-oligomeric states, a dimer and a hexamer.[26,35] The ATP-PRT dimer is composed of two chains that are arranged in a 180° rotational symmetry relative to each other. The interface between two such chains (referred to as the dimer interface) is formed by contacts between residues of domain I of the first chain and domain II of the second chain (and vice versa). The dimer interface buries a surface area of 965 Å^2 to 1203 Å^2 in *M. tuberculosis* ATP-PRT (*Mtu*ATP-PRT)[35] and 1192 Å^2 in *Eco*ATP-PRT.[26] In contrast to the hexamer, the ATP-PRT long form dimer has not been observed crystallographically. Evidence for the existence of the dimer has been presented in the form of kinetic studies,[42] gel-filtration[35,43] and dynamic light scattering experiments.[44]

All structures available for *Mtu*ATP-PRT, *Eco*ATP-PRT and *Mth*ATP-PRT show a homo-hexameric assembly of the enzymes. Herein, three of the described dimers come together side-by-side to create the holo-enzyme. Hexamer-formation is facilitated primarily by contacts between the C-terminal ACT domains. Three ACT domains each come together to create a trimeric assembly, one on each end of the enzyme (Figure 1.7). The ACT domains are arranged with their β sheets facing each other in the centre, burying an individual surface area, for each of the six interfaces, of 790 Å^2 in *Eco*ATP-PRT.[26] The C-terminus is essential for the formation of this interface, contributing almost 70% of the buried ACT domain surface area. It is positioned next to β strand 15 of the adjacent subunit, effectively extending the ACT domain β sheet by a fifth strand, interlocking the trimeric arrangement of the ACT domains.

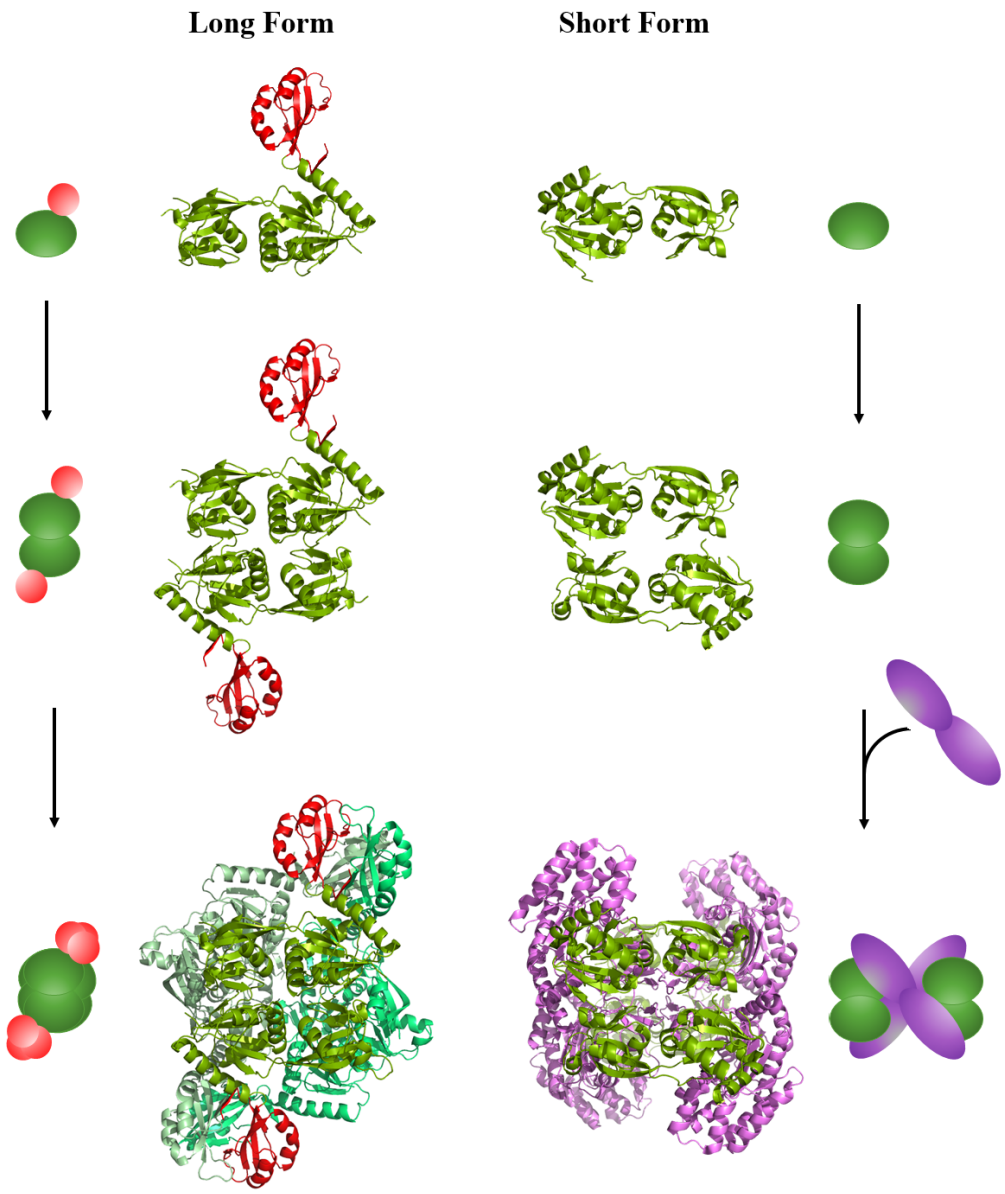


Figure 1.7: Oligomeric assembly of the two ATP-PRT forms. The composition of each form is shown systematically from a single chain to the holo-enzyme. The crystal structures of *Mtu*ATP-PRT (left hand side, long form) and *Lla*ATP-PRT (right hand side, short form) serve as examples. They are displayed in the cartoon representation accompanied by a schematic. LHS: The single chain of long form ATP-PRT enzymes is composed of a catalytic core (green) and a regulatory domain (red). Two chains come together to form a dimer and three dimers associate into a homo-hexamer. RHS: The short form ATP-PRT enzymes also assemble into dimers (green). Together with the regulatory protein HisZ (purple) a hetero-octamer is formed. The schematic is turned by 90 ° compared to the shown complex structure.

Short form ATP-PRT holo-enzymes are assembled in a remarkably different way. Though they form dimers similar to their long form counterparts (Figure 1.7), the holo-enzyme is assembled as a hetero-octamer requiring a second protein, HisZ.[36] HisZ is not only required for the formation of the octamer, but has also been shown to be essential for the activity of *Lla*ATP-PRT.[45] Like its binding partner ATP-PRT, HisZ can form dimers. Two dimers come together in a way that the centre of the resulting complex is formed by a crossed over, "X-shaped", HisZ tetrameric core. The two participating ATP-PRT dimers are bound to either side of this scaffold, completing the arrangement.

1.2.5 Active site of ATP-PRT

The active site of ATP-PRT is positioned in the large cleft between the two domains of the catalytic core. It is constituted by residues of both domains and the two linkers (the β5-β6 and β10-β11 loops). Each active site is connected with a symmetrical counterpart in the dimeric arrangement of the ATP-PRT chains, effectively forming a large cavity on one face of the dimer. In long form ATP-PRT hexamers three of these cavities can be found in the inside of the oligomer, whereas the active sites of short form hetero-octamers face the outside.

Due to the symmetric arrangement of the active sites, it is possible for one chain to contribute amino acid residues across the dimer interface to the neighbouring active site. These cross-interface contributions are likely from the long side chains of the lysine and arginine residues in the β1-α1 and β3-α2 loops, which have been shown to be conserved throughout both ATP-PRT forms.[26,36]

Several ATP-PRT structures have been published that feature active site ligands with complete or partial occupancy. The most prominent of these is the inhibitor molecule AMP bound to the active site of both *Mtu*ATP-PRT[35] (PDB code: 1NH8) and *Eco*ATP-PRT[26] (PDB code: 1H3D). Structural data is also available for the phosphoribosyl-adenine portion of the

product PR-ATP in *Eco*ATP-PRT (PDB code: 1Q1K) and the substrate PRPP in the structure of the short form *Lla*ATP-PRT[36] (PDB code: 1Z7N).

Co-crystallisation of *Eco*ATP-PRT with purified PR-ATP resulted in the crystal structure 1Q1K, although the ligand density presented in this structure only supports the presence of the phosphoribosyl-adenine part of PR-ATP (Figure 1.8A). The binding mode presented for the bound molecule allows for tight interactions of the 5'-phosphoribosyl moiety with residues of the PRPP binding motif of domain II. Hydrogen bonding contacts are formed between the three oxygen atoms and the backbone amides of S172–T176 as well as the side chains of T173 and T176. The ribose oxygens are in hydrogen bonding distance to the conserved glutamate residues E156 and D169. Taken together these interactions are thought to describe the binding of the phosphoribosyl group of PRPP in the long form ATP-PRTs.[26] The adenine group is positioned close to the double peptide linker between domain I and II and is stacked between R16 and L170, which makes up its only interaction with the protein.

PRPP binding to the active site of *Lla*ATP-PRT is observed in all four chains of the crystal structure 1Z7N (Figure 1.8B). Its 5'-phosphoribosyl moiety interacts solely with residues of domain II, primarily of the PRPP binding-motif, analogous to the findings in *Eco*ATP-PRT. Main contacts include the backbone amides of E158–T162 and the side chains of D155, T159 and T162 as well as E142. The orientation of the pyrophosphate is different in each of the four chains but the key interactions are thought to be with S140 as well as to K8', which is contributed across the dimer interface.[36]

The inhibitor molecule AMP has been observed in the *Eco*ATP-PRT crystal structure 1H3D (Figure 1.8C), in a binding mode reminiscent of the one presented for phosphoribosyl-adenine in 1Q1K. AMP effectively mimics (the observed) part of PR-ATP, the only difference being the connectivity of its adenine. Since adenine forms no specific interactions other than the hydrophobic stacking, AMP is perfectly suited to compete for binding in the active site, which explains the competitive character of its inhibition towards both substrates on a molecular level.[26]

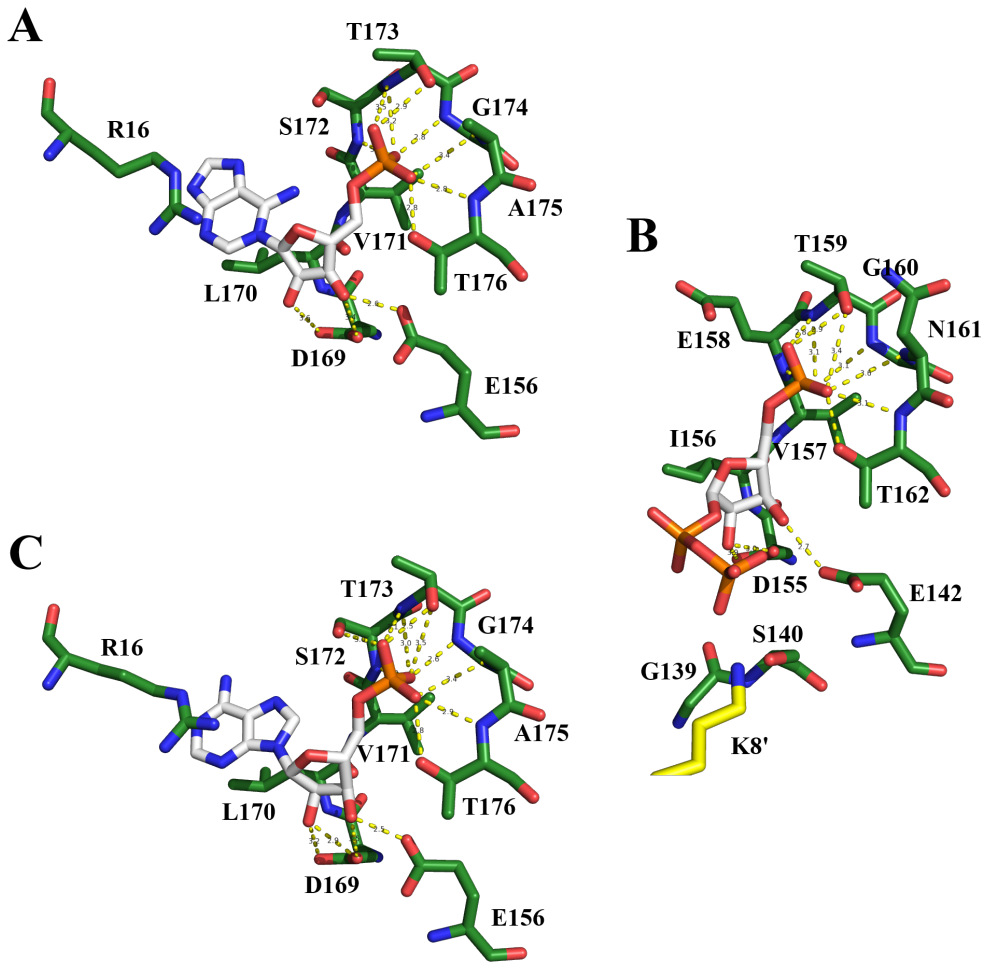


Figure 1.8: ATP-PRT active site ligands. Ligands (white) bound to the active site of ATP-PRT (green) in the stick representation. Yellow dashed lines indicate hydrogen bonding contacts. Hetero atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange). A: Binding mode of phosphoribosyl-adenine in *Eco*ATP-PRT[26] (PDB code: 1Q1K). B: Binding mode of PRPP in *Lla*ATP-PRT[36] (PDB code: 1Z7N). Residue K8' from the adjacent subunit displayed in yellow. C: Binding mode of AMP in *Eco*ATP-PRT (PDB code: 1H3D).

1.2.6 Histidine binding site

The binding of the inhibitor histidine in long form ATP-PRTs takes place at the ACT domain approximately 40 Å away from the active site (regarding a single chain). Histidine binding has been described in the crystal structure of *Mtu*ATP-PRT, which had been co-crystallised in the presence of both histidine and AMP (PDB code: 1NH8). Each chain binds one molecule of histidine. The binding site is situated at the interface of two neighbouring ACT domains, presenting two unequal surfaces. Residues from both chains are recruited for the coordination of histidine. On one side, the carbamino group of histidine forms hydrogen bonding interactions with residues L234–T238, which form a narrow loop, as well as L253 of β14. All of these interactions are formed with the protein backbone except the hydrogen bond between the histidine amino group and the side chain oxygen of the absolutely conserved T238. On the other side, the imidazole moiety is stacked between the side chains of the two aspartate residues D214' and D216' of β12 (Figure 1.9).[35]

The same binding pose for histidine is also found in *Mth*ATP-PRT involving the residues M236–T240, equivalent to the *Mtu*ATP-PRT residues L234–T238, for the binding of the carbamino moiety. In *Mth*ATP-PRT the binding of the imidazole group is achieved by tight interactions with N220', the equivalent of D216'.

The binding of histidine causes a significant conformational change in the *Mtu*ATP-PRT crystal structure. When superimposed on domain I and II of the apo structure (PDB code: 1NH7), the ACT domain of the inhibitor bound structure (PDB code: 1NH8) shows a large rotational displacement of approximately 35°. Due to this motion the holo-enzyme elongates and becomes more compact with a buried subunit-accessible surface area of 3078 Å^2 compared to 2417 Å^2 in the apo form, which is largely attributed to the ACT-ACT interface changes.[35] The histidine bound structure of *Mth*ATP-PRT shows the same conformational change, however such a change was not observed in the AMP and phosphoribosyl-adenine bound crystal structures of *Eco*ATP-PRT. Therefore ATP-PRT long form enzymes adopt either an

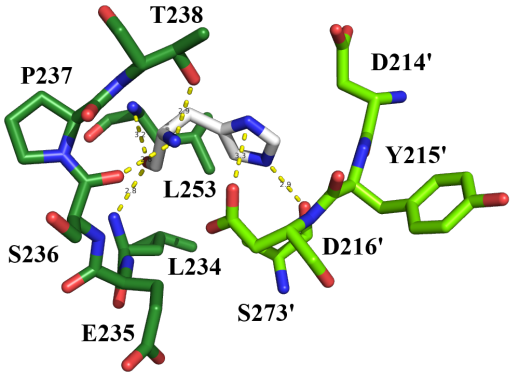


Figure 1.9: Histidine binding site. The histidine binding site of *Mtu*ATP-PRT[35] (PDB code: 1NH8) represented as sticks. Histidine (white) is bound at the interface of two ACT domains (dark green and light green). Yellow dashed lines indicate hydrogen bonding contacts. Hetero atoms are coloured according to element: oxygen (red), nitrogen (blue).

"open" or a "closed" conformation depending on the presence or absence of histidine.[26]

The position of the histidine binding site in the short form ATP-PRT is still debated. Limited crystallographic evidence is given for *Tma*ATP-PRT, in which eight histidine molecules (two per ATP-PRT monomer) are placed in two equivalent sites at the interfaces formed with HisZ.[37] In contrast to this observation stands the mutational analysis of *Lla*ATP-PRT that suggests the involvement of residues in a putative histidine binding site located solely in HisZ.[36] Whether one or both hypotheses are correct is still unclear.

1.2.7 Evolution

The evolutionary relationship of the separate units of both long and short form ATP-PRTs have been examined, both to each other, and to other protein families. The catalytic core unit of domain I and II has been compared to the family of the monomeric periplasmic binding proteins,[35] which are directly involved in the uptake of nutrients, such as sulphate and phosphate as well as amino acids, in bacteria (reducing the need for *de novo* synthesis). Periplasmic binding proteins, such as the *E. coli* HisJ,[46] are composed of two

domains that share a similar fold with the catalytic core of ATP-PRT with an RMSD of 2.57 Å between the two *E. coli* proteins, suggesting a common ancestry.[36] The high specificity of the PRPP binding loop for negatively charged ions could therefore be the evolutionary relic from an ancestral protein involved in sulphate and phosphate binding. Low sequence similarity and the different function and motion[35,47] of the two enzyme families on the other hand stand in clear contrast to this theory.

The regulatory ACT domain of the ATP-PRT long form shares high structural similarity with members of the P_{II} protein superfamily, one of the most widely distributed signalling proteins in nature.[48] P_{II} proteins act as sensory devices for the metabolic state of the cell, a feature that they share with the ATP-PRT regulatory domain. Furthermore, P_{II} proteins adopt the characteristic $\beta\alpha\beta\beta\alpha\beta$ fold with a C-terminal β strand and assemble into functional trimers near identical to the ACT domain of ATP-PRTs. The main differences between the two structures are the elongated C-terminus and the long T-loop of P_{II}, which is inserted between the first and second canonical $\beta\alpha\beta$ unit. In ATP-PRT ACT domains both structural elements are significantly shorter. Interestingly, both protein families bind their respective ligand (ATP in P_{II}, histidine in ATP-PRT) at the interface between two neighbouring (ACT) domains. Consequently it has been proposed that an ancestral P_{II} or P_{II}-like domain was recruited to create the long form of ATP-PRT.[49]

A second gene product, HisZ, is essential for ATP-PRT short form structure[37] and function.[45] HisZ is a class II aminoacyl-tRNA synthetase paralogue that lacks the anti-codon binding domain of aminoacyl-tRNA synthetases, but possesses an α-helical insertion domain, which is essential for the formation of the hetero-octameric complex of short form ATP-PRTs. Phylogenetic studies have shown that HisZ is only encoded in bacterial histidine operons that also code for a short ATP-PRT gene. It has also been found that short and long form ATP-PRTs generally do not co-exist, the Geobacter clade forming the only known exception,[50] dividing the histidine synthesising species into two heterogeneous groups.[51] The short form ATP-PRT, which is

only found in bacteria, is believed to be the evolutionary older form, based on its similarity with aminoacyl-tRNA synthetase, linking back to the early stages of life. The long form ATP-PRT is believed to have evolved later in either bacteria or archaea and was subsequently transferred into eukarya.[52]

In summary, all ATP-PRT enzymes are likely to have evolved from a common ancestor and then diverged during early evolution into the two forms present today (Figure 1.10). From the hypothetical precursor, that is potentially related to periplasmic binding proteins, two co-evolutionary pathways originate. The first used an aminoacyl-tRNA synthetase paralogue as regulatory subunit for the formation of a hetero-complex (ATP-PRT short form), while the second recruited a P_{II}-like protein as a covalently linked regulatory domain (ATP-PRT long form).

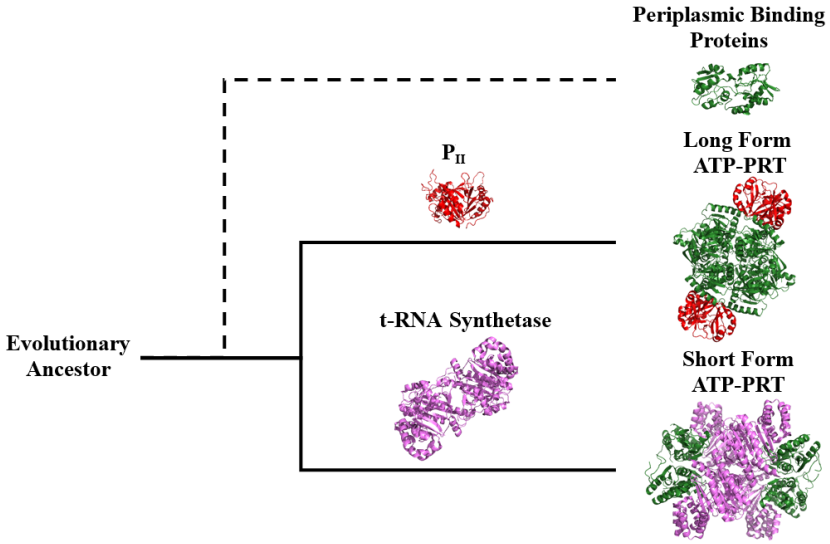


Figure 1.10: Evolution of ATP-PRT. Overview over the theoretical evolutionary pathways presented for ATP-PRT. Equivalent structural elements have identical colours.

1.3 Enzyme regulation

The biochemistry of the cell is a vast network of enzymatic pathways by which metabolites are generated, converted and degraded. Situated at key junctions of these pathways are regulatory enzymes, which control the metabolic flux. They are able to sense the metabolic needs of the cell and adjust their catalytic activity accordingly. The diverse mechanisms that control these enzymes ensure the efficient use of the available nutrient and energy pools and consequently the survival of the cell.[53]

ATP-PRT is one of these key enzymes. Catalysing the reaction between two important cellular metabolites, ATP and PRPP, it is situated at the beginning of the energetically expensive histidine biosynthetic pathway.[14] Cells that exhibit unregulated histidine biosynthesis have been shown to dissipate up to 2.5 % of their metabolic energy.[54] It is therefore essential to regulate ATP-PRT, which is achieved by a multi-layered mechanism, mainly concentrated on the modulation of the enzyme activity and to a lesser extent the expression level of the protein.

1.3.1 Transcriptional regulation

The transcription level of bacterial operons can be regulated by attenuation.[55] In amino acid biosynthetic operons attenuation involves a transcribed leader region in the operon preceding the first structural gene. The essential part of this leader region is a segment coding for a short polypeptide, which is rich in the amino acid that is the end product of the corresponding biosynthetic pathway. This segment acts as a sensory device for the abundance of the relevant charged tRNAs and only allows the transcription of the operon to proceed under amino acid starvation conditions.[56]

The histidine operon, encoding for all genes of the histidine biosynthetic pathway, is controlled by attenuation and does not co-exist with a repressor-operator system as seen for other amino acid operons, therefore making it

the only form of genetic level regulation for the abundance of ATP-PRT, and consequently histidine, in the bacterial cell.[12]

1.3.2 Active site inhibition

One of the most common mechanisms for enzyme inhibition is the reversible binding of a small molecule, the inhibitor, to (the active site of) an enzyme, preventing the substrate binding by mere physical exclusion. This competitive inhibition can be overcome by a high substrate to inhibitor ratio.[57]

Individually, AMP and ADP inhibit ATP-PRT with a clear competitive character towards both substrates, ATP and PRPP.[18] In the case of AMP, this is strongly supported by the observed binding pose in the active site of *Mtu*ATP-PRT and *Eco*ATP-PRT (Figure 1.8C). The energy level of the cell can generally be described by the ratio of ATP to AMP (and ADP). Since ATP is the substrate and AMP and ADP inhibitors of ATP-PRT, the enzyme is directly sensing the energy state of the cell and forced to adjust activity accordingly. During starvation the ATP/AMP ratio will strongly decrease, inhibiting the enzyme and stopping the metabolic flux into the energetically expensive histidine *de novo* synthesis.[21]

Additionally, ATP-PRT has been reported to be inhibited by the products of the biosynthetic direction reaction, PR-ATP and PP. This logically results from the fact that the ATP-PRT catalysed reaction is an equilibrium.[17,28] Physiologically this would result in an arrest of ATP-PRT catalysis in the biosynthetic direction only if the histidine pathway is blocked at the second (irreversible) step, catalysed by PR-ATP pyrophosphohydrolase.

The effectiveness of AMP inhibition alone is based solely on the general energy state of the cell, but a defining feature of ATP-PRT is the coupling of this regulation mechanism with the inhibition by the allosteric feedback inhibitor histidine, resulting in an overall synergistic effect.[8,18]

1.3.3 Allostery

The term 'allostery' or 'allosteric control' describes the regulatory control of an enzyme by the binding of a small effector molecule at a site other than the protein active site. It is derived from the two Greek words '$\alpha\lambda\lambda o\varsigma$' (other) and '$\sigma\tau\varepsilon\rho\varepsilon o\varsigma$' (solid object) and is meant to express the 'difference' (allo-) in the (stereo-) specificity of the effector and substrate binding sites.[58] Allostery can affect the catalytic properties of an enzyme positively (activation) or negatively (inhibition). It can be differentiated between homotropic and heterotropic allostery. In homotropic allostery, typically known as cooperativity, the substrate of an enzyme can also act as effector at the same time. In heterotropic allostery the effector is not a substrate for the enzyme. Generally the inhibition by pathway end products is achieved by heterotropic allostery.

Principally, allostery can affect an enzyme's conformation or dynamics. Due to their intrinsic conformational flexibility, proteins are able to sample a variety of different conformer populations. Effector binding interferes with the equilibrium between those populations and alters the distribution of conformer species, ultimately resulting in changes of the active site geometry and consequently the catalytic efficiency.[59] A change between structurally distinguishable conformations is not always necessary. Changes in protein dynamics can also contribute decisively to allosteric control.[60,61]

The evolution of allosteric models

The first allosteric effects were discovered at the beginning of the 20th century. The observation of the sigmoidal behaviour of oxygen binding to haemoglobin[62] soon led to the formulation of the Hill equation,[63] a fundamental instrument of enzyme kinetics. The next key observations were made in the 1950s, when metabolic pathway end products (particularly amino acids) were found to inhibit their own production by affecting an early enzyme in the pathway (to which they are not a substrate), creating a negative

feedback regulation.[64,65] In 1963 the term 'allosteric' was introduced to describe these systems[66] and in 1965 the concerted Monod-Wyman-Changeux (MWC) model[67] was postulated, which describes allostery as the transition of (oligomeric) enzymes between two possible conformations. Soon afterwards the sequential Koshland-Nemethy-Filmer (KNF) model[68] was formulated to address the ability of individual subunits in an oligomer to change conformation independently. Most allosteric systems can be explained in broad terms by these two models, such as the classic model proteins haemoglobin and aspartate transcarbamylase.[69] However, in recent years a third model has arisen from the need to explain the behaviour of the enzyme porphobilinogen synthase, which exists in an equilibrium between quaternary structure isoforms with different kinetic properties. This model has been named the "morpheein" model after the words 'protein' and 'to morph' (to change shape), describing the ability of a morpheein to assemble into structurally and functionally distinct oligomers.[70] The key features of the three models are summarised below (Figure 1.11).

In the concerted MWC model multimeric enzymes are described to exist in an equilibrium between two distinct symmetric states, R (relaxed) and T (tense). The simplified assumption is made that all subunits exist in the same state at any given time. To maintain the molecular symmetry, all subunits must transform 'in concert' from one conformation into the other. The allosteric effector shifts the position of the equilibrium by preferentially binding and stabilizing one of the states (T). According to this model, even low effector concentrations will have a large effect of the enzyme's overall conformation and function.

The sequential KNF model adopts the basic principle of the two equilibrium states and augments it by the proposal that the change in one subunit of a protein does not necessarily induce the same change in any other subunit. Based on the induced fit hypothesis[71] each subunit is able to change conformation independently upon effector binding, resulting in mixed conformational, asymmetric intermediate species.

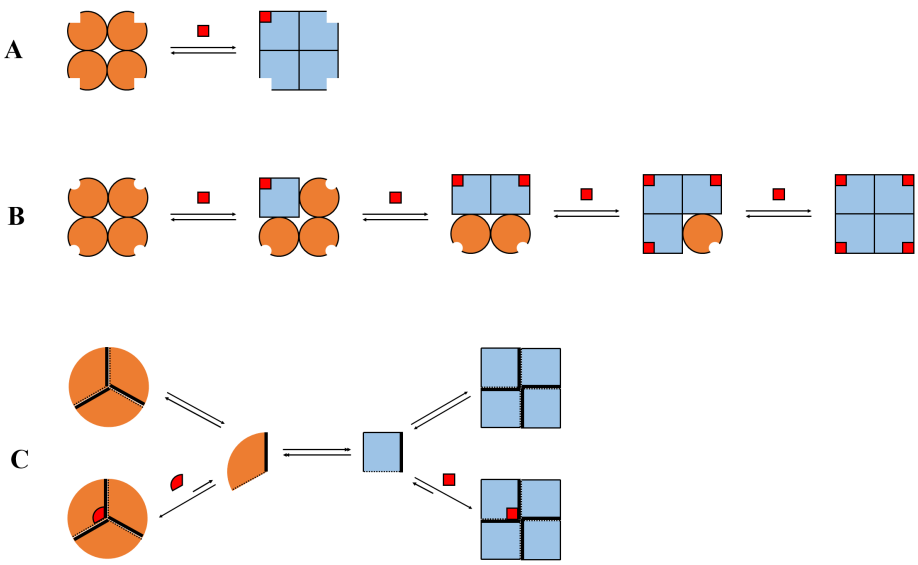


Figure 1.11: Models of allostery. Two dimensional schematic representation of the three allosteric models. Different conformations of the subunits are indicated by colour (orange and blue). Effector molecules are highlighted in red. A: The concerted MWC model.[67] A symmetric multimeric protein can exist in an equilibrium between two conformational states that differ in their affinity for the effector. B: The sequential KNF model.[68] Each protein subunit can change its conformation independently. The allosteric binding site adjusts to accommodate the effector ("induced fit"). C: The morpheein model.[70] A homo-oligomeric protein exists in an equilibrium between different monomeric conformations, each dictating a different, distinct oligomeric assembly. Oligomers must assemble their interfaces symmetrically (fat line to dashed line). Effector molecules can shift the equilibrium towards certain oligomers.

In contrast to the MWC and KNF models, where the oligomeric state of the enzyme is fixed, the central postulate of the morpheein model is the necessity for quaternary structure change. In this model, homo-oligomeric enzymes can readily associate and dissociate under physiologic conditions. The dissociated (monomeric) state exists in an equilibrium between different conformational states (morpheeins) that each dictate the formation of a different oligomeric assembly with defined multiplicity and different functionality. According to the model, the assembly of the morpheeins must preserve the symmetrical contacts between the subunits in each oligomer. The morpheein equilibrium can be influenced through the binding and stabilisation

by morpheein-specific effectors, drawing the equilibrium towards a specific oligomer.

Regardless of these models, allostery, in its general sense, can also be applied to monomeric proteins.[72,73]

Allosteric control of ATP-PRT by the feedback inhibitor histidine

Both forms of ATP-PRT have been shown to be under allosteric control by the pathway end product histidine by classic feedback inhibition.[13,37] But due to their vastly different oligomeric architecture, the molecular mechanisms that underlie their regulation appear to be different.

In long form ATP-PRT the feedback inhibition by histidine is directly associated with a conformational change between an open and a closed form, which has been directly observed in the crystal structures of *Mtu*ATP-PRT (Figure 1.12). The change is based on a large twist of the ACT domain, relative to the domains I and II. Histidine binding occurs at the interface of two neighbouring ACT domains, suggesting that direct interactions with histidine, including a well ordered hydrogen bonding network, are causing ATP-PRT to adopt the closed conformation.[35]

Only the binding of histidine results in the described conformational changes. The crystal structure of *Eco*ATP-PRT in the presence of AMP (PDB code: 1H3D) shows the same overall conformation as described for the apo form of *Mtu*ATP-PRT (PDB code: 1NH7), underlining that this allosteric control mechanism is solely performed by the feedback inhibitor histidine. Small differences in the positioning and the presented interfaces between the ACT domains of AMP bound *Eco*ATP-PRT and the apo *Mtu*ATP-PRT structures have been described though, presenting the ACT-ACT interface of *Eco*ATP-PRT in a marginally tighter position. These differences have been attributed to the effect of AMP binding to the active site, which could explain the synergistic inhibition of AMP and histidine described for ATP-PRTs.[26]

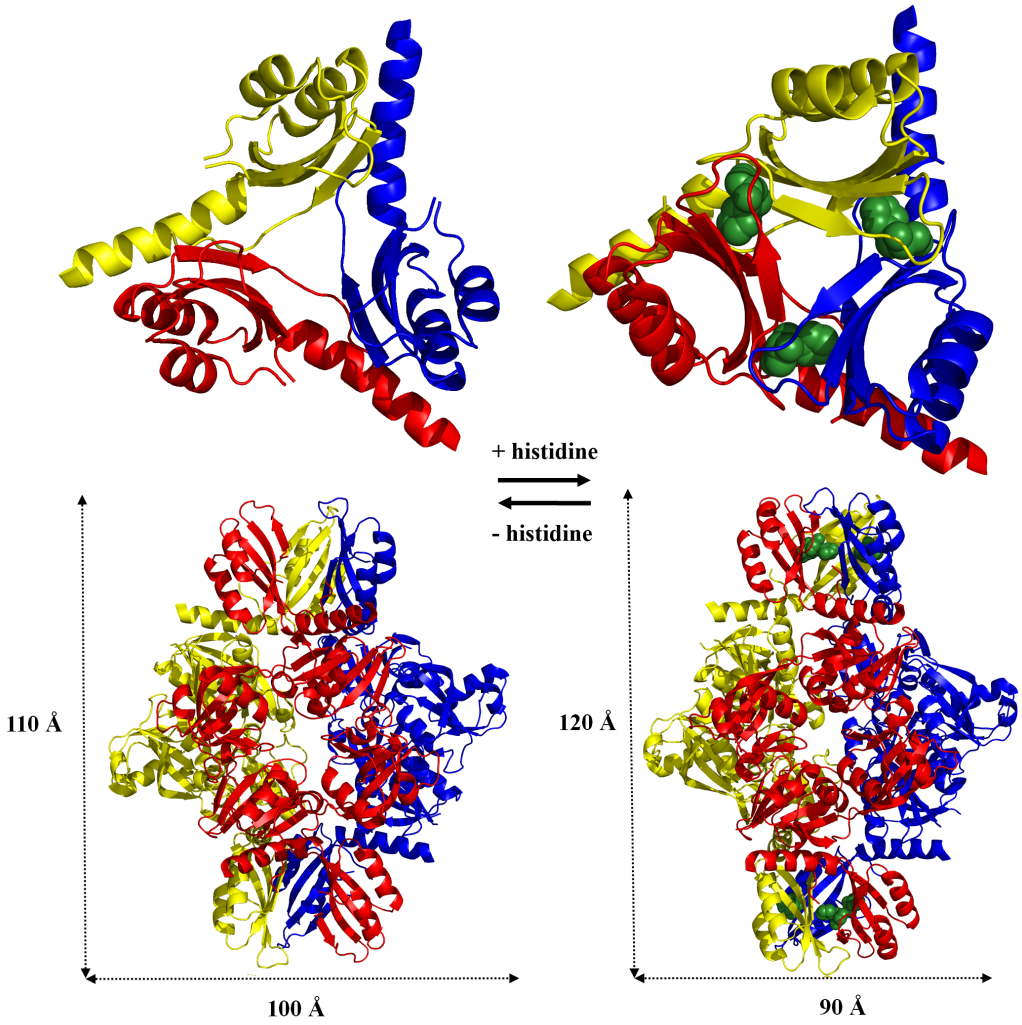


Figure 1.12: Allosteric mechanism of histidine feedback inhibition. Comparison of the open (left hand side) and closed (right hand side) crystal structures of *Mtu*ATP-PRT (PDB codes: 1NH7 and 1NH8)[35] in a cartoon representation. Dimeric units are displayed in the same colour. Histidine is shown as green spheres. Top: View on the trimeric arrangement of the ACT domains. Upon binding of histidine the ACT domains twist, bringing their β sheets closer together in the centre. Bottom: The open and closed hexamer labeled with their approximate molecular dimensions in two axes. In the absence of histidine the holo-enzyme is significantly wider and shorter.

According to the current models, long form ATP-PRT enzymes exist in an equilibrium between active dimers and inactive hexamers. Although all long form ATP-PRT crystal structures show hexameric arrangements, it

is assumed that the hexamer is not a permanent state. The catalytic dimer is believed to be the solely active species, whereas the hexameric state is only adopted in solution upon histidine binding or at high protein concentration.[74] The binding of histidine at the ACT domain interfaces allows for strong interactions and leads to hexamer formation, while the dissociation of histidine also promotes the enzyme's dissociation into dimers. The apo structure of *Mtu*ATP-PRT (PDB code: 1NH7) and the two *Eco*ATP-PRT crystal structures (PDB codes: 1Q1K and 1H3D) are therefore thought to present loosely attached dimers forced together by crystal packing. This opinion is justified by their relatively weak ACT domain interactions. The change of oligomeric state, from dimer to hexamer, induced by histidine binding, resulting in either altered topology or reduced access to the active site, is consequently considered to be the main allosteric mechanism of long form ATP-PRT enzymes.[5,75]

Comparative analysis of the short form ATP-PRT structures suggested a potential role of the ATP-PRT dimer interface in the allosteric mechanism. Upon binding of histidine, this interface is altered, causing the removal of potentially catalytically relevant residues from the nearby active sites.[36] Hereby of special interest are conserved lysine and arginine residues, as they are thought to create cross-dimer contacts and help in the binding or activation of PRPP. K8 and K50 of *Lla*ATP-PRT change their positioning in respect to the neighbouring active site. Mutagenesis of both residues resulted in reduced catalytic properties of *Lla*ATP-PRT, which supports their importance for catalysis.[27] These findings have also been related to the structures of the long form enzymes *Mtu*ATP-PRT and *Eco*ATP-PRT and interpreted as a general mode of allosteric regulation in ATP-PRT enzymes.[36]

1.4 Goals of this thesis

The main goal of this study is to gain a deeper understanding of the molecular basis for the regulation of ATP-PRT enzymes, with a focus on the long form and its characteristic regulatory domain. To obtain a wider perspective, a new member of this enzyme group was chosen and its comprehensive characterisation is presented, including detailed analysis of its crystal structure. To extend the current knowledge of the ATP-PRT ligand interactions and inhibitory synergy, the binding properties of this enzyme are also studied. Furthermore the regulatory and catalytic role of the C-terminal regulatory ACT domain are tested by separating it from the catalytic core structure and studying the individual components in comparison to the wild type. Mutational analysis is reported and reveals the involvement of crucial amino acid residues in a potential allosteric communication pathway between the remote inhibitor binding site and the active site. Finally, all gathered information is discussed in regards to the regulatory mechanisms of long form ATP-PRT enzymes.

The detailed objectives of this thesis were to:

- clone, purify and characterise a novel member of the ATP-PRT long form from *Campylobacter jejuni* (*Cje*ATP-PRT);
- investigate the oligomeric state of *Cje*ATP-PRT in solution;
- produce and solve the crystal structures of substrate and inhibitor bound *Cje*ATP-PRT;
- generate a C-terminal truncation mutant, by removing the entire regulatory domain of *Cje*ATP-PRT, and investigate its enzymatic properties and evolutionary implications;
- investigate a potential communication pathway between the binding sites of *Cje*ATP-PRT using single site mutants.

Chapter 2

Adenosine triphosphate phosphoribosyltransferase from *Campylobacter jejuni*

2.1 Introduction

Characterisation of ATP-PRT began in the late 1950s and the pioneering studies were carried out on the long form enzyme *Sen*ATP-PRT.[13,16] The kinetic characterisation of this enzyme led to a detailed understanding of its kinetic properties and complex inhibition.[18,28] The sedimentation behaviour of *Eco*ATP-PRT provided first insights into the oligomery of the ATP-PRT long form.[74,76] Decades later, the crystal structures of *Mtu*ATP-PRT[35] and *Eco*ATP-PRT[26] were solved, which, for the first time, enabled the exploration of the relationship between structure and function of long form ATP-PRTs. Based on this knowledge, the mutational analysis of *Corynebacterium glutamicum* ATP-PRT (*Cgl*ATP-PRT) resulted in the identification of crucial residues for the interactions with histidine.[77]

Following on from these results, this study is aimed to consolidate a general view on the overall properties and structure of the ATP-PRT long

form, to establish some of the finer details of its regulation and to explore the evolution of the enzyme. To widen the perspective, a new member of this enzyme sub-family was cloned and characterised. *C. jejuni* was chosen as the enzyme source.

C. jejuni is classified as a Gram-negative proteobacterium of the ε clade[78] and has a spiral rod-shaped appearance. It naturally colonises the digestive tracts of many bird species[79] but is also found in a variety of other animal faeces including cattle, flies, and kangaroo. Despite its appearance in other animals, *C. jejuni* causes serious infections in humans and is the main cause of bacterial gastroenteritis in the developed world, although it is rarely fatal.[80,81] Given the increasing number of antimicrobial agent-resistant strains worldwide,[80] characterisation of the enzymes of the histidine biosynthetic pathway might also provide the opportunity for new antibiotic treatments in the future.

The genome of *C. jejuni* subsp. *jejuni* NCTC 11168, first sequenced in 2000,[82] contains a single gene encoding for a putative ATP-PRT long form as identified by sequence similarity. Cloning of this gene, as well as the purification and characterisation of the encoded enzyme, *Cje*ATP-PRT, are presented in this chapter.

2.2 Multiple sequence alignment and phylogenetic analysis of the ATP-PRT long form

To obtain a systematic overview of the ATP-PRT long form sequences available, a large primary sequence alignment was created including 307 characterised and putative ATP-PRT long form enzymes from the KEGG database.[83] Sequences were chosen on the basis of length ($\geq$ 270 amino acids) and genus (not more than two species of the same genus). Therefore over-representation of some taxonomic groups (e.g. enterobacteria) is due to the higher number of complete or partially sequenced genomes available in these taxa. Plastidic

signal peptides were ignored for the purpose of this alignment. A visual summary in the form of a phylogram is given in Figure 2.1. Protein sequences with high similarity were identified and grouped in taxonomic clusters. A consensus sequence was established for each cluster and subsequent alignment of these consensus sequences was performed to produce a more representative result (Figure 2.2).

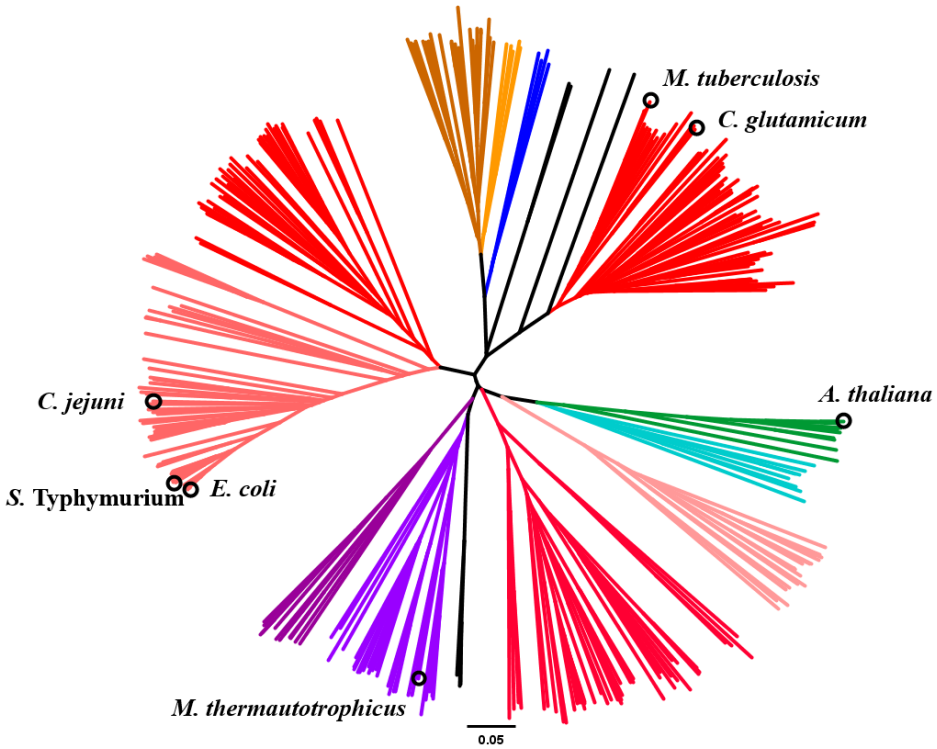


Figure 2.1: Phylogram of selected ATP-PRT long form sequences. Alignment was performed using ClustalOmega[84] and the resulting data is displayed as a phylogram generated using FigTree.[85] Scale of genetic distance is displayed below. The colour of the clades corresponds to systematic grouping: plants (green), cyanobacteria and chloroflexi (cyan), bacteria (red), proteobacteria (pink), archaea (purple), fungi (orange and brown), other eukaryotes (blue). Positions of characterised enzymes are indicated by black circles.

Following from the phylogenetic analysis, the ATP-PRT long form type can be sub-divided into four main groups with various subgroups. The internal group sequence identity spans from 21 to 99 % whereas inter-group identity is not higher than 38 %. Sequences encoding functionally char-

acterised or putative long form ATP-PRTs are found in a large variety of organisms including archaea, bacteria, fungi, plants, and lower eukaryotes, a ubiquitous distribution with animals being the only exception. On a general view the tree spreads according to the current opinions on systematics and evolution.[86] This supports the theory that the histidine biosynthetic pathway along with many other amino acid biosynthetic pathways, as a basic functionality, was established very early during evolutionary history.[52]

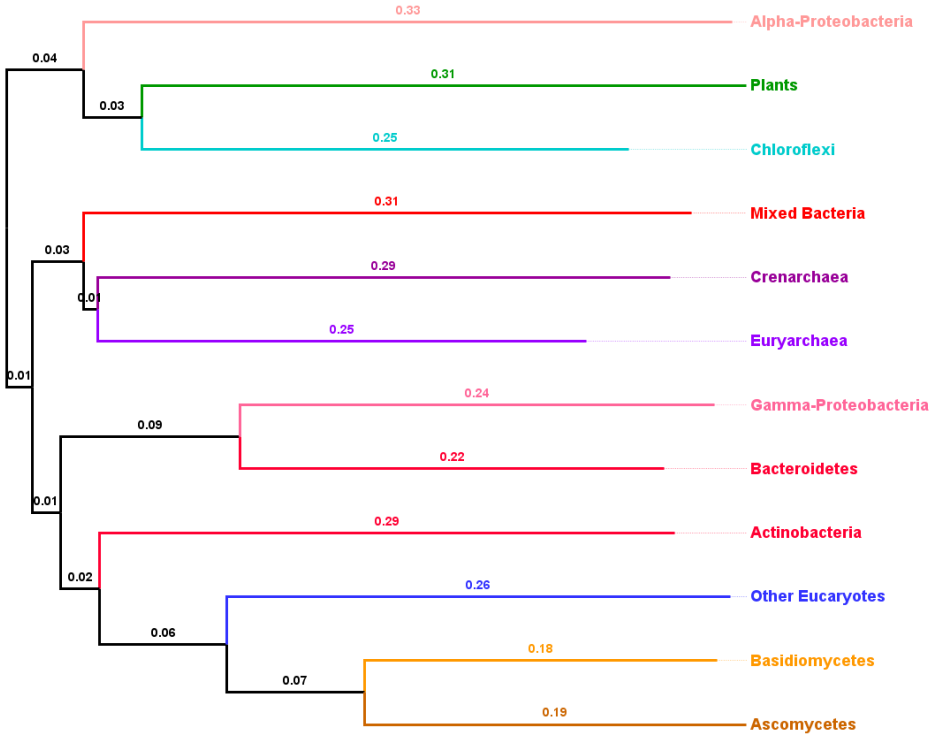


Figure 2.2: Phylogram of ATP-PRT long form consensus sequences. Alignment was performed using ClustalOmega,[84] resulting data is represented as phylogram generated using FigTree.[85] Genetic distance is displayed above the branches. Clades were coloured according to Figure 2.1.

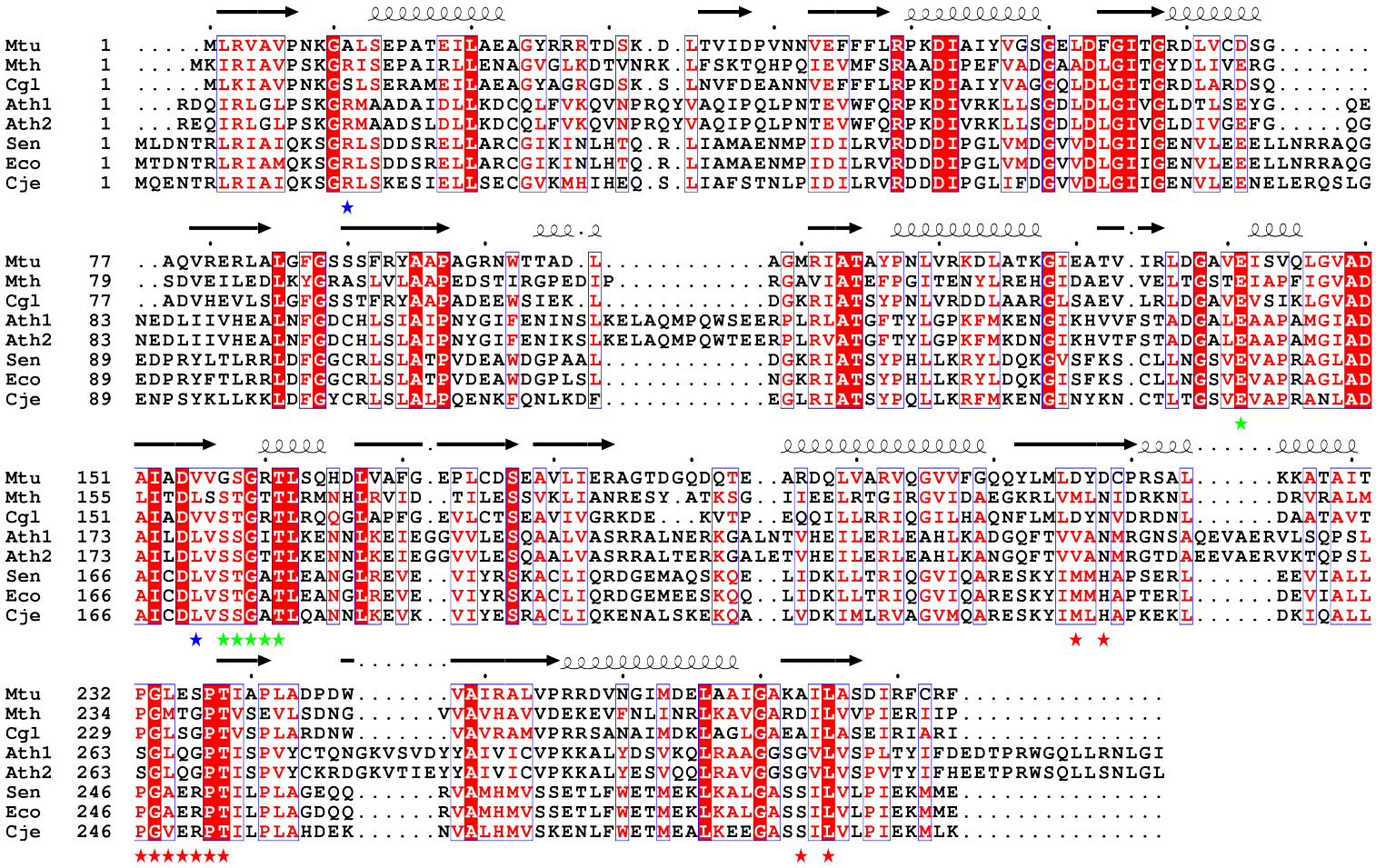
One of the most obvious findings is the similarity between the protein sequences from plants and the mixed group of cyanobacteria and chloroflexi (green non-sulphur bacteria), an observation that can be interpreted in two ways. Either both groups have been under a very similar evolutionary pressure, regarding their histidine biosynthesis, or plants have acquired the *hisG* gene from a member of the cyanobacteria/chloroflexi group via lateral gene

transfer early in evolution, most likely during an endosymbiotic event. According to the endosymbiotic theory, cyanobacteria are thought to be the endosymbionts, which would be in accordance with the hypothesis of the lateral gene transfer. This is further supported by the plastidic location of the histidine biosynthetic enzymes in plants.[9,87]

To allow further comparison and identification of key residues for small molecule binding and catalysis, a primary sequence alignment of *Cje*ATP-PRT with all structurally and/or biochemically characterised long form enzymes was performed. Following the previous alignments (Figure 2.1 and 2.2) that grouped *Cje*ATP-PRT with the γ proteobacterial ATP-PRTs, a somewhat unique positioning, the proteins with the highest sequence similarity to *Cje*ATP-PRT were *Eco*ATP-PRT and *Sen*ATP-PRT with overall identities of 66 % and 65 % respectively (Figure 2.3).

The ATP-PRT long form sequences show an overall low conservation. Only 35 residues are conserved between the compared enzymes with the majority being clustered in two areas: the α2-β4 fold of domain I and the PRPP binding motif (β9-α7) of domain II, the PRT signature fold (Figure 1.4). The histidine binding loop in domain III is also found to be composed of generally conserved residues. The sequence of this short motif is strictly PGXXXPT, except for in the *A. thaliana* enzymes where the first proline residue is replaced by a serine.

Figure 2.3 *(following page)*: Multiple sequence alignment of characterised ATP-PRT long form enzymes. The alignment was created with ClustalOmega[84] and edited using ESPript.[88] Secondary structure elements of *Mtu*ATP-PRT are depicted above as bolt arrows (β strands) and spirals (α helices). Residues are coloured according to conservation: absolutely conserved (white, red box), partially conserved (red) and not conserved (black). Residues known to interact with histidine (red), 5'-phospho-ribose (green) and adenine (blue) groups are highlighted below. Organism three letter code: *M. tuberculosis* (Mtu), *M. thermautotrophicus* (Mth), *C. glutamicum* (Cgl), *A. thaliana* (Ath), *S. enterica* subsp. *enterica* Typhimurium (Sen), *E. coli* (Eco) and *C. jejuni* (Cje).


Mtu 1 .....MLRVAVPNKGALSEPATEILAEAGYRRRTDSK.D.LTVIDPVNNVEFFFLRPKDIAIYVGSGELDFGITGRDLVCDSG.....
Mth 1 ....MKIRIAVPSKGRISEPAIRLLENAGVGLKDTVNRK.LFSKTQHPQIEVMFSRAADIPEFVADGAADLGITGYDLIVERG.....
Cgl 1 .....MLKIAVPNKGSLSERAMEILAEAGYAGRGDSK.S.LNVFDEANNVEFFFLRPKDIAIYVAGGQLDLGITGRDLARDSQ.....
Ath1 1 ...RDQIRLGLPSKGRMAADAIDLLKDCQLFVKQVNPRQYVAQIPQLPNTEVWFQRPKDIVRKLLSGDLDLGIVGLDTLSEYG.....QE
Ath2 1 ...REQIRLGLPSKGRMAADSLDLLKDCQLFVKQVNPRQYVAQIPQLPNTEVWFQRPKDIVRKLLSGDLDLGIVGLDIVGEFG.....QG
Sen 1 MLDNTRLRIAIQKSGRLSDDSRELLARCGIKINLHTQ.R.LIAMAENMPIDILRVRDDDIPGLVMDGVVDLGIIGENVLEEELLNRRAQG
Eco 1 MTDNTRLRIAMQKSGRLSDDSRELLARCGIKINLHTQ.R.LIAMAENMPIDILRVRDDDIPGLVMDGVVDLGIIGENVLEEELLNRRAQG
Cje 1 MQENTRLRIAIQKSGRLSKESIELLSECGVKMHIHEQ.S.LIAFSTNLPIDILRVRDDDIPGLIFDGVVDLGIIGENVLEENELERQSLG
Mtu 77 ..AQVRERALGFGSSSFRYAAPAGRNWTTAD.L.....AGMRIATAYPNLVRKDLATKGIEATV.IRLDGAVEISVQLGVAD
Mth 79 ..SDVEILEDLKYGRASLVLAAPEDSTIRGPEDIP.....RGAVIATEFPGITENYLREHGIDAEV.VELTGSTEIAPFIGVAD
Cgl 77 ..ADVHEVLSLGFGSSTFRYAAPADEEWSIEK.L.....DGKRIATSYPNLVRDDLAARGLSAEV.LRLDGAVEVSIKLGVAD
Ath1 83 NEDLIIVHEALNFGDCHLSIAIPNYGIFENINSLKELAQMPQWSEERPLRLATGFTYLGPKFMKENGIKHVVFSTADGALEAAPAMGIAD
Ath2 83 NEDLIIVHEALNFGDCHLSLAIPNYGIFENIKSLKELAQMPQWTEERPLRVATGFTYLGPKFMKDNGIKHVTFSTADGALEAAPAMGIAD
Sen 89 EDPRYLTLRRLDFGGCRLSLATPVDEAWDGPAAL.....DGKRIATSYPHLLKRYLDQKGVSFKS.CLLNGSVEVAPRAGLAD
Eco 89 EDPRYFTLRRLDFGGCRLSLATPVDEAWDGPLSL.....NGKRIATSYPHLLKRYLDQKGISFKS.CLLNGSVEVAPRAGLAD
Cje 89 ENPSYKLLKKLDFGYCRLSLALPQENKFQNLKDF.....EGLRIATSYPQLLKRFMKENGINYKN.CTLTGSVEVAPRANLAD
Mtu 151 AIADVVGSGRTLSQHDLVAFG.EPLCDSEAVLIERAGTDGQDQTE..ARDQLVARVQGVVFGQQYLMLDYDCPRSAL.....KKATAIT
Mth 155 LITDLSSTGTTLRMNHLRVID..TILESSVKLIANRESY.ATKSG..IIEELRTGIRGVIDAEGKRLVMLNIDRKNL.....DRVRALM
Cgl 151 AIADVVSTGRTLRQQGLAPFG.EVLCTSEAVIVGRKDE...KVTP..EQQILLRRIQGILHAQNFLMLDYNVDRDNL.....DAATAVT
Ath1 173 AILDLVSSGITLKENNLKEIEGGVVLESQAALVASRRALNERKGALNTVHEILERLEAHLKADGQFTVVANMRGNSAQEVAERVLSQPSL
Ath2 173 AILDLVSSGTTLKENNLKEIEGGVVLESQAALVASRRALTERKGALETVHEILERLEAHLKANGQFTVVANMRGTDAEEVAERVKTQPSL
Sen 166 AICDLVSTGATLEANGLREVE..VIYRSKACLIQRDGEMAQSKQE..LIDKLLTRIQGVIQARESKYIMMHAPSERL.....EEVIALL
Eco 166 AICDLVSTGATLEANGLREVE..VIYRSKACLIQRDGEMEESKQQ..LIDKLLTRIQGVIQARESKYIMMHAPTERL.....DEVIALL
Cje 166 AICDLVSSGATLQANNLKEVK..VIYESRACLIQKENALSKEKQA..LVDKIMLRVAGVMQARESKYIMLHAPKEKL.....DKIQALL
Mtu 232 PGLESPTIAPLADPDW.....VAIRALVPRRDVNGIMDELAAIGAKAILASDIRFCRF.....
Mth 234 PGMTGPTVSEVLSDNG.....VVAVHAVVDEKEVFNLINRLKAVGARDILVVPIERIIP.....
Cgl 229 PGLSGPTVSPLARDNW.....VAVRAMVPRRSANAIMDKLAGLGAEAILASEIRIARI.....
Ath1 263 SGLQGPTISPVYCTQNGKVSVDYYAIVICVPKKALYDSVKQLRAAGGSGVLVSPLTYIFDEDTPRWGQLLRNLGI
Ath2 263 SGLQGPTISPVYCKRDGKVTIEYYAIVICVPKKALYESVQQLRAVGGSGVLVSPVTYIFHEETPRWSQLLSNLGL
Sen 246 PGAERPTILPLAGEQQ.....RVAMHMVSSETLFWETMEKLKALGASSILVLPIEKMME.....
Eco 246 PGAERPTILPLAGDQQ.....RVAMHMVSSETLFWETMEKLKALGASSILVLPIEKMME.....
Cje 246 PGVERPTILPLAHDEK.....NVALHMVSKENLFWETMEALKEEGASSILVLPIEKMLK.....

2.3 Cloning and expression

Successful cloning of the *C. jejuni hisG*, locus tag Cj1597, was achieved using the Invitrogen Gateway® Cloning system as described in section 6.1.16. With the *C. jejuni* subsp. *jejuni* NCTC 11168 genomic DNA, supplied by ATCC® (ATCC strain identification: 700819), as the template, a linear DNA construct of the estimated size (approximately 950bp), verified by agarose gel electrophoresis, was generated in three consecutive rounds of PCR. This construct included a tobacco etch virus (TEV) protease cleavage site immediately upstream of the *hisG* gene. The linear fragment was inserted into the shuttle vector pDONR™ 221 using the Gateway® BP Clonase™ and transformed into TOP10 cells for selection, followed by plasmid propagation. The purified plasmid was sequence verified and used in the LR Clonase™ reaction, which transferred the *C. jejuni hisG* gene into the expression vector pDEST™17 with an N-terminal poly-histidine tag (His-tag).

The resulting construct was used for heterologous expression in BL21* (DE3) cells. SDS-PAGE of the induced crude lysate showed successful overexpression, with one major band just below the 40 kDa molecular weight marker corresponding to the estimated size of tagged *Cje*ATP-PRT (approximately 37.2 kDa), but situated mostly (>90 %) in the insoluble fraction. In order to produce more overexpressed protein in the soluble fraction, the pDEST™17 *Cje*ATP-PRT construct was transformed into the Chaperone 3 cell line. This resulted in an increase in the amount of soluble *Cje*ATP-PRT. ATP-PRT activity was detected in the soluble crude using the standard ATP-PRT activity assay (section 6.1.33) and this final expression construct was thus used for all protein purification.

2.4 Purification

A purification strategy was devised based on the removable N-terminal His-tag. The enzyme was isolated from the crude lysate by immobilised metal

affinity chromatography (IMAC) using Ni^{2+}-charged nitrilotriacetic acid (Ni-NTA), then desalted and freed of the His-tag by TEV protease treatment. The tag-free enzyme was then to be reverse purified by a second IMAC step and finally purified and desalted by size exclusion chromatography (SEC). However, two major problems were encountered with this protocol. Firstly, TEV protease was less efficient at removing the His-tag than expected. A series of different incubation conditions, designed to test the influence of TEV protease concentration, temperature, stirring speed, and time, were attempted to improve the cleavage efficiency. An initial incubation at 37°C and agitation at 180 rpm, followed by a period without shaking or stirring at 4°C for 3 days gave the best improvement in tag cleavage efficiency. However, even under these optimised conditions a small portion of tagged enzyme remained. The removal of the tag had no adverse effects on enzyme activity (Table 2.1). Secondly, *Cje*ATP-PRT showed an unexpected strong intrinsic affinity to the column. This was noticeable both in the initial IMAC, where the elution occurred relatively late (at 250–300 mM imidazole concentration), but mostly the second IMAC passage when the tag-free *Cje*ATP-PRT bound the column instead of eluting in the loading buffer (containing 20 mM imidazole) as expected. This led to changes in the purification protocol.

Table 2.1: Purification steps for *Cje*ATP-PRT. Enzyme activity was measured at each purification step using the standard assay described in section 6.1.33. Protein concentrations were determined according to section 6.1.29 to calculate the specific activity. Recovery is a measure of the loss of the total activity of the whole preparation.

Purification step	Total activity [U]	Specific activity [U mg^{-1}]	Recovery [%]
Soluble crude	334 ± 48	0.11 ± 0.02	100 ± 14
IMAC eluate	246 ± 8	1.70 ± 0.06	74 ± 2
TEV protease treated	237 ± 8	1.64 ± 0.05	71 ± 2
SEC eluate	171 ± 8	2.44 ± 0.12	51 ± 2

To overcome the problems caused by the high affinity of the protein to the Ni-NTA resin, a different column material was selected for the IMAC purification steps. By using a HiTrap™ TALON® crude column the elution of the protein was shifted towards the earlier part of the elution gradient suggesting a lower affinity. This also reduced the imidazole concentration of the eluate significantly, which positively influenced the protein stability and rendered the desalting step unnecessary. Since the protein was found to bind to the IMAC column in the absence of the His-tag, the second IMAC passage was removed from the purification protocol. During the experiments

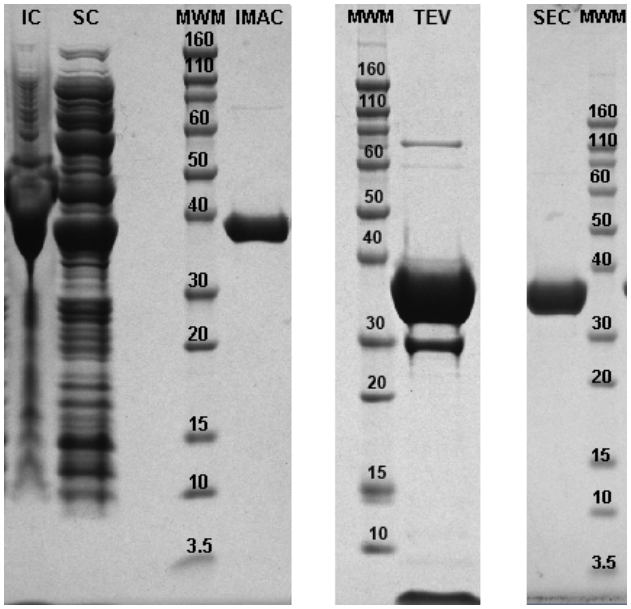


Figure 2.4: *Cje*ATP-PRT purification steps. Samples were taken after each purification step and assessed for purity by reducing SDS-PAGE in 10 % Bis-Tris gels using MES or MOPS buffer. Denaturation was performed by incubation at 95°C for 2 min before sample loading. The main species in all preparations was equivalent to the monomeric size of *Cje*ATP-PRT, which consistently appeared on the gel at an apparent higher molecular weight (tagged: 37.2 kDa, untagged: 33.7 kDa). The minor impurity carried through all purification steps ($\approx$ 65 kDa) is most likely dimeric *Cje*ATP-PRT, appearing perhaps due to insufficient denaturation time. The band at 30 kDa corresponds to TEV protease while the low molecular weight band ($\approx$ 4 kDa) is the cleaved tag. Lane labels correspond to: MWM - Molecular weight marker, IC - Insoluble crude, SC - Soluble crude, IMAC - Eluate from IMAC, TEV - TEV protease treated, SEC - Eluate from SEC.

to improve the purification protocol it was observed that the overall yield was dependent on the length of time the cell pellet had been stored. Generally a fresh, unfrozen pellet gave a much higher yield than a frozen pellet and the older the frozen pellet the lower the yield.

The final purification procedure was the following. The cells were lysed immediately following harvest, when possible, and the *Cje*ATP-PRT was isolated from the soluble fraction by IMAC. The eluate underwent TEV protease treatment and was purified and desalted by the final SEC step. The pure protein was either kept at 4°C for immediate use or concentrated and flash frozen. All steps were performed as described in detail in chapter 6. Figure 2.4 shows the purity of the *Cje*ATP-PRT after each step and Table 2.1 lists the corresponding activities. With this protocol an overall yield of 15–20 mg pure *Cje*ATP-PRT per litre of liquid culture was achieved.

2.5 Determination of molecular weight

The molecular mass of the purified *Cje*ATP-PRT was determined to be 33695.2 g mol^{-1} by electrospray ionisation mass spectrometry. This value is in close agreement with the theoretical value of 33696.2 g mol^{-1} calculated for the 300 amino acid sequence, which includes the complete native *Cje*ATP-PRT open reading frame and a single N-terminal glycine residue left over from the truncated tag. A peak corresponding to a tagged version of the enzyme was not detected.

2.6 Kinetic properties

Detailed kinetic characterisation of the pure *Cje*ATP-PRT was performed as it is essential to understand the functionality of the enzyme. The ATP-PRT reaction is the reversible conversion of ATP and PRPP into PR-ATP and PP. It is possible to follow the production of PR-ATP directly by the increase of

absorbance at 290 nm due to the increased aromaticity of PR-ATP compared to ATP. Using one of the substrates in excess (usually ATP) this increase is also directly proportional to the consumption of the limited substrate. All measurements of kinetic activity that are described here were performed using a standard assay (section 6.1.33) following the comprehensive kinetic analysis of the *Sen*ATP-PRT enzyme,[17,18,28] with minor adjustments that are described in the following sections.

2.6.1 Pyrophosphatase

The reaction catalysed by the ATP-PRT is an equilibrium (Figure 1.2). In the conditions of the assay, this equilibrium lies at approximately 50–55 % in the biosynthetic direction with a determined K_{eq} of 0.9. To ensure the complete turnover of all substrate and eliminate the possible effects on the initial rate measurements, the assay was supplemented with an auxiliary enzyme, inorganic pyrophosphatase (EC: 3.6.1.1). Pyrophosphatase catalyses the irreversible hydrolysis of PP into phosphate removing all PP, which drives the ATP-PRT reaction to completion.

For this study the *E. coli* inorganic pyrophosphatase wild type (*Eco*PPase) was cloned using the Gateway® Cloning procedure, expressed, and purified using the introduced His-tag (section 6.2.2). The activity of the purified *Eco*PPase was tested by adding it to the reaction after the equilibrium had been established, which resulted in a second absorbance rise corresponding to the complete turnover of the limiting substrate (in most cases PRPP).

It was also possible to follow the *Cje*ATP-PRT reaction in the reverse direction, where PR-ATP and PP are consumed to produce ATP and PRPP. When *Cje*ATP-PRT was mixed only with PR-ATP and PP a decrease in absorbance was observed. After it had reached the equilibrium point the addition of *Eco*PPase reversed the direction and led to an increase in absorbance corresponding to the production of PR-ATP (Appendix A).

2.6.2 General behaviour

*Cje*ATP-PRT retained catalytic activity over a long period of time when stored at 4°C. When stored at room temperature the first significant decrease in enzyme activity (8 %) was seen after 48 hours, whereas rapid freeze/thaw cycles had a strong negative effect with more than 60 % activity loss after the second cycle (Figure 2.5). Together with the finding that fresh cell pellets produce the highest protein yield it can be assumed that freezing adversely affects the enzyme both *in vitro* and *in vivo*.

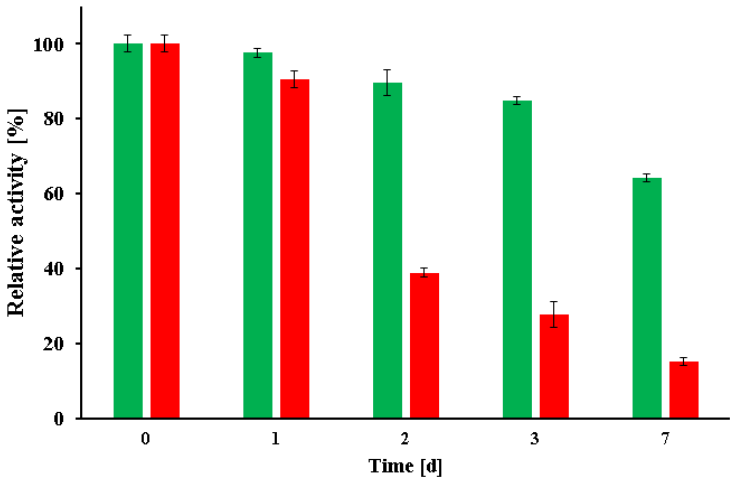


Figure 2.5: Activity decay over time. Relative activity was measured under standard assay condition using 21.2 μg mL^{-1} *Cje*ATP-PRT. The activity of freshly purified enzyme was monitored over the course of a week while stored at room temperature (green) or -80°C being frozen and thawed once between each measurement (red).

As for other kinetically characterised ATP-PRT enzymes, *Cje*ATP-PRT activity is dependent on the concentration of Mg^{2+} ions in the assay solution, as well as the concentration of salt (NaCl and KCl) used.[28,89] To determine the optimal assay conditions these parameters were investigated in more detail. The results are shown in Figure 2.6. *Cje*ATP-PRT showed increased activity over a concentration range of 20–250 mM for NaCl and 20–300 for KCl. The maximum activity measured in the presence of salt was about 40 % higher than when no salt was added to the assay. Roughly 120 % activity was measured with 100 mM NaCl or 50 mM KCl. A mixture of

50 mM of each salt had approximately the same effect, but was surpassed by combinations of 100 mM NaCl or KCl with 50–100 mM of the other salt. Above an overall salt concentration of 500–600 mM the activity slowly decreased to a measured minimum of 41 % at 600 mM KCl and 500 mM NaCl. The minor advantage of potassium over sodium is likely due to their different position in the Hofmeister series. K^+ is less chaotropic than Na^+, which might in turn increase the stability of the enzyme. As a result of this analysis, the standard assay used in this work contained 100 mM NaCl and 50 mM KCl.

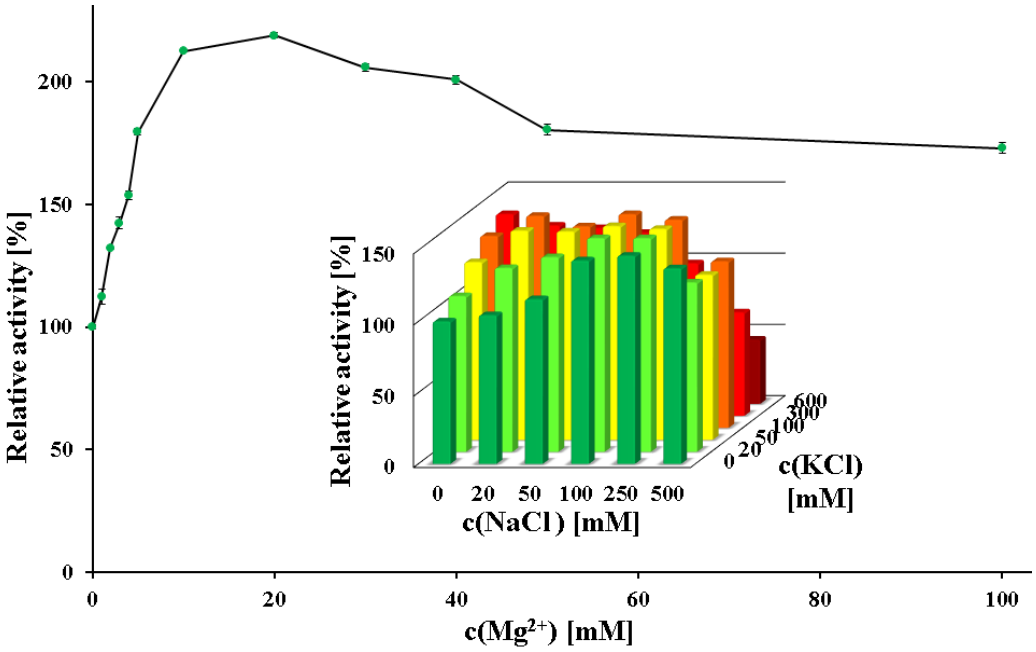


Figure 2.6: Influence of Na^+, K^+ and Mg^{2+} concentration on the activity of *Cje*ATP-PRT. Specific activity was measured under the conditions described in section 6.2.3 and normalised against the activity in the absence of metal(s). The magnesium dependence is shown as a scatter graph with lines directly connecting the single measurements (green). Inset: three dimensional bar graph showing the relative activity of *Cje*ATP-PRT in the presence of 36 Na^+ and K^+ combinations. With increasing K^+ concentration, the bar colour changes gradually from green to red.

The optimal concentration of Mg^{2+} for *Cje*ATP-PRT in the assay was determined in the presence of 100 mM NaCl and 50 mM KCl. After an initial sharp rise maximal activity was reached at a concentration between 10 and 20

mM. Further increasing the Mg^{2+} concentration, the relative activity slowly levelled off to approximately 170 % of the initial activity at 0 mM Mg^{2+}. The optimal Mg^{2+} concentration for *Cje*ATP-PRT was higher than the reported value for *Sen*ATP-PRT[28] (9–10 mM) in the presence of 5 mM ATP.

It is possible that the labile substrate PRPP is subject to hydrolysis under assay conditions forming PP and ribose-5-phosphate (R5P). To confirm that it was stable, PRPP (120 μM) was incubated for 1 h in a complete assay mix, without the second substrate ATP, before the ATP-PRT reaction was initiated with the addition of ATP. A second assay was prepared lacking only the *Cje*ATP-PRT and treated in the same manner. Both initial rate and total turnover were measured and compared to an identical assay that had been initiated without incubation time. Approximately 9 % of the PRPP had been lost after 1 h in the reaction lacking ATP and over 18 % in the reaction lacking *Cje*ATP-PRT, showing an unexpected, stabilising effect of *Cje*ATP-PRT on PRPP under assay conditions.

2.6.3 Kinetic parameters

Michaelis-Menten constants (K_M) for the substrates ATP and PRPP were determined from non-linear fitting (2.7). Both K_M values were found to be in the micro-molar range with the K_M(ATP) being about 6-fold larger. The *Cje*ATP-PRT showed consistently low turn-over rates. The obtained Michaelis-Menten constants compare well to previously published data on *Sen*ATP-PRT[28] but are 3-fold and 6-fold lower than data reported for *Mtu*ATP-PRT[89] and *Ath*ATP-PRT[87] respectively. A summary is given in Table 2.2.

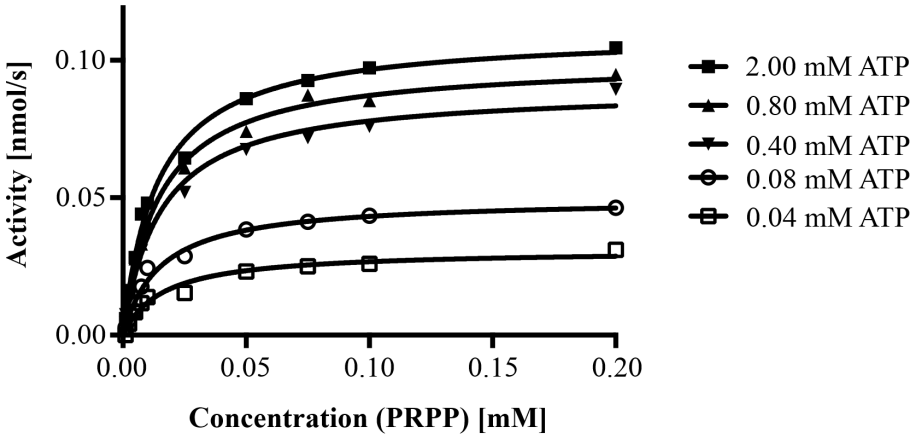


Figure 2.7: Determination of *Cje*ATP-PRT kinetic parameters. Initial velocity data for *Cje*ATP-PRT was collected according to sections 6.1.33 using the substrate concentrations described in section 6.2.3. Non-linear fitting was performed using GraphPad Prism version 6.00 (GraphPad Software).

Table 2.2: Comparison of ATP-PRT kinetic constants. Kinetic constants of *Cje*ATP-PRT were determined under the conditions described in section 6.2.3 and compared to values reported for *Sen*ATP-PRT,[28] *Ath*ATP-PRT[87] and *Mtu*ATP-PRT.[89]

Organism	K_M(ATP) [μM]	K_M(PRPP) [μM]	k_{cat} [s^{-1}]
C. jejuni	97 ± 10	15 ± 1	1.8 ± 0.1
S. Typhimurium[28]	110	15	n.d.
A. thaliana[87]	600	130	n.d.
M. tuberculosis[89]	263 ± 63	49 ± 6	n.d.

2.6.4 Binding mechanism

Enzymatic reactions with multiple substrates (and products) can be classified according to the sequence of substrate binding. Two main mechanisms can be differentiated. In the sequential mechanism the enzyme binds substrate molecules sequentially prior to turn over. The products are then released from the active site one by one. Binding and release may either follow a strict order (ordered sequential) or not (random sequential). In the alternative ping-pong mechanism the enzyme only binds and turns over one substrate at a time, alternating between states with different substrate affinities. The double reciprocal plots of the activity sets measured for an enzyme at constant concentrations of one of the substrates either intersect (sequential mechanism) or they are parallel (ping-pong mechanism).

*Cje*ATP-PRT was tested for its binding mechanism in the biosynthetic direction reaction by measuring a grid of non-saturating substrate concentrations (6.2.3). The resulting data shows clear evidence for a sequential binding mechanism as the theoretical extensions of the plots converge (Figure 2.8).

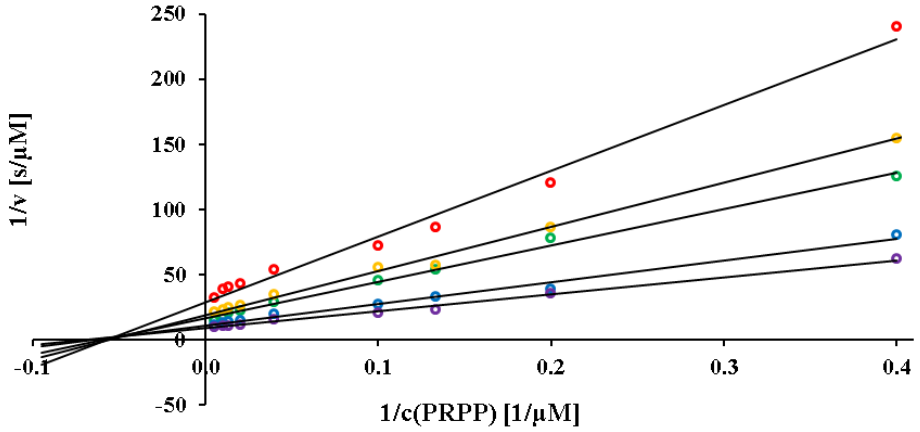


Figure 2.8: Sequential binding mechanism. Lineweaver-Burk plot of the initial rates measured for the biosynthetic direction reaction of *Cje*ATP-PRT. Rates were measured at PRPP concentrations between 1 and 200 μM and the plotted ATP concentrations were: 40 μM (red), 80 μM (orange), 200 μM (green), 400 μM (blue) and 2.0 mM (purple). The lines merge close to the X-axis, representative of a sequential binding mechanism.

This is consistent with previous findings for ATP-PRT from other species.[28] Further kinetic analysis concerning the type (ordered vs. random) have not been made as neither the product PR-ATP nor a mimic of PR-ATP were available (in sufficient quantities) for product inhibition studies.

2.6.5 Inhibition

In accordance with previous investigations on long form ATP-PRT enzymes, *Cje*ATP-PRT was found to be inhibited by AMP and histidine (Table 2.3). Inhibition studies for these two inhibitors were performed against both substrates and the resulting data was fitted with the equations for all four possible inhibition modes. The AMP inhibition data against both substrates was consistent with a competitive model yielding values K_i(AMP) around 300–400 μM, indicating a relatively weak effect under assay conditions. A non-competitive model produced the best fit for the histidine inhibition data against both substrates with K_i values of around 40–50 μM (Appendix C).

Table 2.3: Inhibition constants of AMP and histidine. K_i values for AMP and histidine on *Cje*ATP-PRT were determined as described in section 6.2.3.

Inhibitor	Inhibition mode	K_{iATP} [μM]	K_{iPRPP} [μM]
AMP	Competitive	356 ± 22	382 ± 34
Histidine	Non-competitive	44 ± 7	40 ± 1

While both histidine and AMP are inhibitors individually, their effect is increased when both are present. This synergistic effect was clearly observable at low concentrations when both inhibitors only marginally affected the activity. Addition of both inhibitors to the assay resulted in a significant drop in activity that was larger than the theoretical additive value of both inhibitors separately (Figure 2.9A). The synergy is also evident by plotting the reciprocal initial velocities in the presence of different histidine concentrations against the AMP concentration (Figure 2.9B). The resulting graphs

are curved and not linear as would have been expected if histidine and AMP were acting independently of each other.

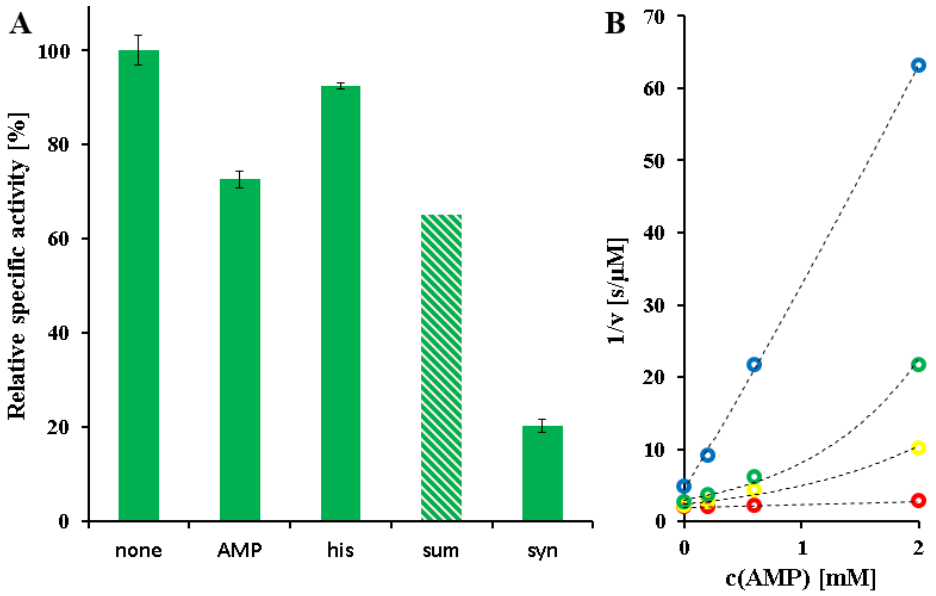


Figure 2.9: Synergistic inhibition by AMP and histidine. Activity was measured according to section 6.1.33 using 2 mM ATP and 0.3 mM PRPP. The influence of the two inhibitors on each other is depicted as A: Bar graphs showing the inhibitory response to the relative activity in the presence of no inhibitor (none), 2 mM AMP (AMP), 20 μM histidine (his) and 2 mM AMP plus 20 μM histidine (syn). The activity of the uninhibited enzyme reduced by both individual inhibitor responses is given as a theoretical comparison (sum). B: Reciprocal plot of initial velocity against AMP concentration at four different histidine concentrations: 50 μM (blue), 30 μM (green), 20 μM (yellow) and 0 μM (red). Exponential fits to the data points are displayed as black dotted lines.

2.7 Ligand binding

The binding interactions of the ATP-PRT with its natural ligands are largely unknown. The structural information available features only the inhibitors histidine and AMP,[26,35] but there are at least eight natural ligands known to bind to ATP-PRT, each of two substrates and products, Mg^{2+}, and three inhibitors. This creates a complex network of potential interplay between

multiple ligands and the protein. Having kinetically characterised *Cje*ATP-PRT, the effects of ligands on the protein were investigated using calorimetry. The stoichiometry, binding affinity and thermodynamics of ligand interactions with *Cje*ATP-PRT were studied. This was mostly achieved through the application of ITC, which had not previously been reported for ATP-PRT enzymes.

2.7.1 Isothermal titration calorimetry

To investigate the binding interactions of *Cje*ATP-PRT with its ligands, ITC was used to measure the dissociation constants of four natural ligands alone and in the presence of a second ligand. All measurements were performed as described in section 6.2.4 (Figure 2.10).

Binding of the substrates ATP and PRPP to *Cje*ATP-PRT was measured individually, with K_d values, calculated from the fit to a one-site model, of 158 ± 13 for ATP and 122 ± 8 for PRPP, which are respectively 1.6-fold and 6-fold higher than the determined apparent K_M values in the presence of saturating concentrations of the corresponding second substrate. Contrary to expectations, the enzyme bound PRPP with a dissociation constant in the same order of magnitude as that determined for ATP (Table 2.4). The large difference to the apparent K_M therefore indicates that PRPP binding is strongly affected by the presence of ATP. Unfortunately it was not possible to study the direct influence of the substrates to one another as turnover would have occurred in the cell when titrating one substrate into the cell holding *Cje*ATP-PRT saturated with the second substrate.

Figure 2.10 *(following page)*: Ligand binding of *Cje*ATP-PRT. ITC binding curves were measured under the conditions described in section 6.1.39 and 6.2.4, subtracted by their corresponding heat of dilution data series and fitted using OriginLab® version 7.0. Single plots are showing the binding of A: ATP, B: PRPP, C: AMP, D: Histidine, E: ATP in the presence of histidine, F: AMP in the presence of ATP, G: AMP in the presence of histidine, H: Histidine in the presence of AMP.

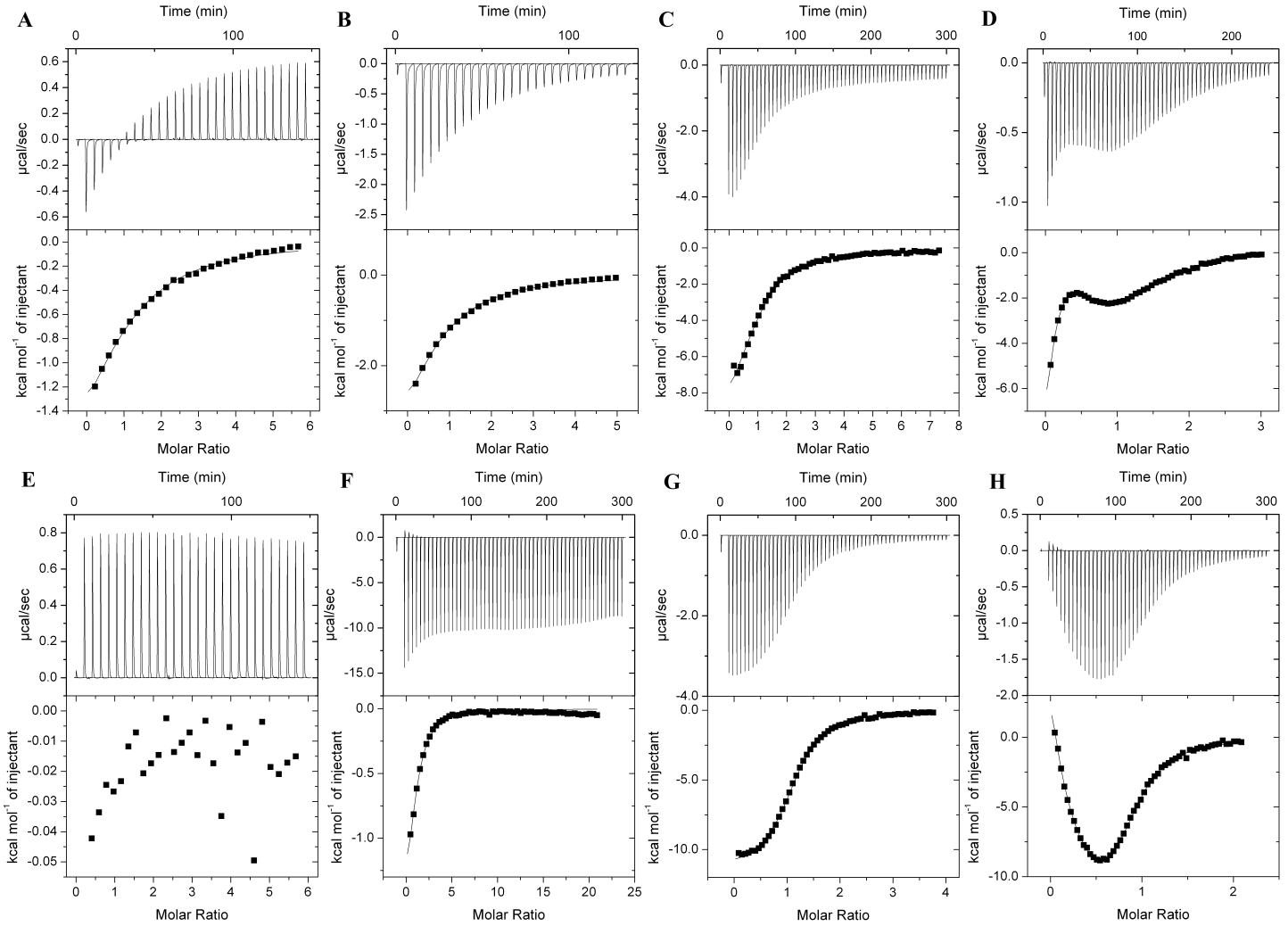

A
B
C
D
E
F
G
H
Time (min)
µcal/sec
kcal mol^{-1} of injectant
Molar Ratio

Table 2.4: Dissociation constants. Dissociation constants of *Cje*ATP-PRT ligands were determined by ITC according to section 6.2.4. A stoichiometry for the histidine titration cannot be determined (ND) due to the type of fit used.

Ligand	ATP	PRPP	AMP	Histidine
K_d [μM]	158 ± 13	122 ± 8	51 ± 4	19.3 ± 0.2, 14.5 ± 0.2, 36 ± 1, 50 ± 1, 47 ± 1
Stoichiometry	1.07 ± 0.08	0.72 ± 0.05	1.00 ± 0.04	ND

The investigation of the inhibitor binding properties of *Cje*ATP-PRT showed two distinctly different results. AMP binding fits a standard one site model but the resulting K_d is 6–8-fold lower than the K_i values measured kinetically. Histidine on the other hand consistently exhibited a complex binding behaviour that was only able to be fitted by applying a sequential model with ≥ 4 binding events. This may suggest a strong cooperative effect of the individual binding sites of the enzyme complex, a conformational change of the enzyme, or both simultaneously. The K_d values determined from the fit are close to the kinetic inhibition constants, confirming the strength of this interaction.

To investigate the competitive relationship of AMP and ATP, AMP was titrated into the enzyme in the presence of 1 mM ATP. The data was still consistent with a one site model. The resulting change in the K_d of AMP, about 3-fold higher than in the absence of ATP, confirmed that both molecules are competing for the same binding site (Table 2.5). This explained why the kinetically determined K_i value measured in the presence of 2 mM ATP (around 400 μM), was a lot higher than the K_d value measured in the absence of ATP.

Strong evidence for the synergistic effect of AMP and histidine was found by titrating AMP into *Cje*ATP-PRT completely saturated with histidine, or vice versa. Not only did the affinity of the enzyme for each inhibitor increase significantly in the presence of the other (Table 2.5), but the appearance of the histidine titration curve was altered notably. The ini-

Table 2.5: Interference of *Cje*ATP-PRT ligands. Interference of *Cje*ATP-PRT ligands was measured using ITC according to section 6.2.4. Dissociation constants and stoichiometry were determined by fitting the data to a single site model (AMP) or a sequential model (histidine). Abbreviations used are: no binding observed (NB), cannot be determined (ND).

Titrated ligand	ATP	AMP	AMP	Histidine
Ligand in cell	Histidine	ATP	Histidine	AMP
K_d [μM]	NB	83 ± 8	8.3 ± 0.2	4.8 ± 0.7, 2.3 ± 0.1
Stoichiometry	ND	1.1 ± 0.1	1.14 ± 0.01	0.9 ± 0.2

tial strong exothermic signal disappeared, leaving the distinctive cooperative shape of two binding events. This demonstrates clearly that the presence of AMP alters the binding interactions for histidine. It also confirms that a bi-directional signal transduction exists between the active site, where AMP binds, and the allosteric histidine binding site.

2.7.2 Influence on thermal stability

To further investigate the influence of ligand binding to the *Cje*ATP-PRT, thermal stability measurements were performed using differential scanning calorimetry (DSC) and differential scanning fluorimetry (DSF) (section 6.1). Limited success was achieved with the DSF experiments as an initial high fluorescence signal was observed in all samples, with the exception of the sample containing AMP and histidine (Figure 2.11). A high initial signal is usually indicative of premature protein denaturation. In the case of *Cje*ATP-PRT, the signal decreased with increasing temperature prior to displaying a typical melting behaviour.

Since the known ATP-PRT ligands are generally very polar and the enzyme showed no sign of destabilisation prior to the experiment, the initial high fluorescence was interpreted as nonspecific binding of the dye to the

enzyme complex. Unfortunately information on possible specific site interactions is lacking as the molecular structure of the SYPRO® Orange dye is proprietary. The high background reading may have altered the properties of the melting event. Hence DSC experiments were performed to confirm the DSF observations.

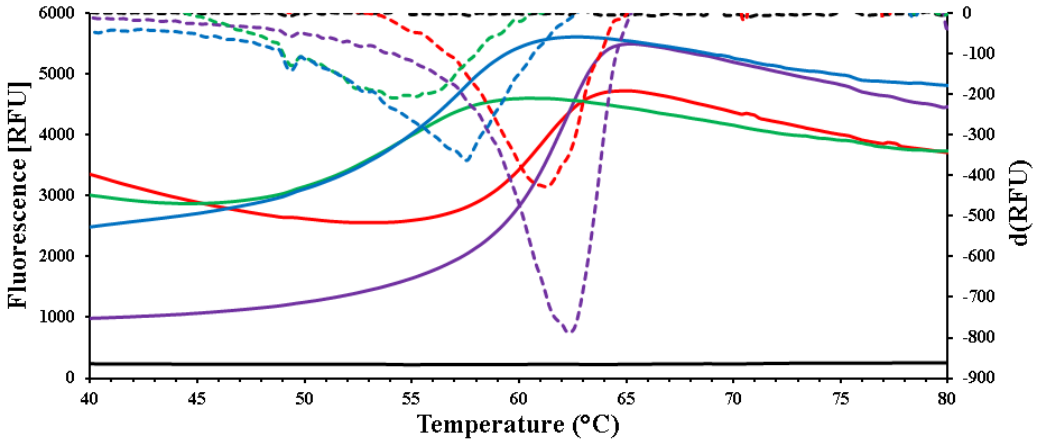


Figure 2.11: DSF analysis of *Cje*ATP-PRT. Melting curves (solid lines) of 0.1 mg mL $^{-1}$ enzyme samples were determined according to section 6.1.37 in the presence of 5 mM $MgCl_2$ and no ligand (green), 2 mM AMP (red), 1 mM histidine (blue) and a mix of 2 mM AMP and 1 mM histidine (purple). A water blank is given in black. The first derivatives (dotted lines) show the inflection points of the melting curves, which represent the observed average T_m values.

According to the DSF results, pure, untreated enzyme is stable up to 57.5 °C in the presence of 5 mM $MgCl_2$ prior to thermal denaturation. The addition of AMP shifted the melting temperature (T_m) by 3 °C upwards, an effect that was enhanced even further with a mixture of AMP and histidine seen as a shift of 5 °C, whereas the addition of histidine alone only altered the melting temperature by 2 °C. The presence of the substrate ATP on the other hand had no influence on the melting temperature. Most interestingly, the AMP plus histidine sample showed low initial dye binding leading to the conclusion that the binding of AMP and histidine must lead to a change of enzyme shape or dynamics rendering the enzyme more rigid and less susceptible to thermal denaturation.

Very similar results were acquired from the DSC experiments (Figure 2.12), in which the stabilising effects of AMP and histidine were confirmed. The thermal shifts measured by DSC are in close agreement with the DSF values, althought the T_m values measured were consistently 3° higher, which is likely to be caused by the faster temperature ramping speed of the DCS experiments (1 °C/min vs. 0.2 °C/20 s). A summary of all resulting T_m values is given in Table 2.6.

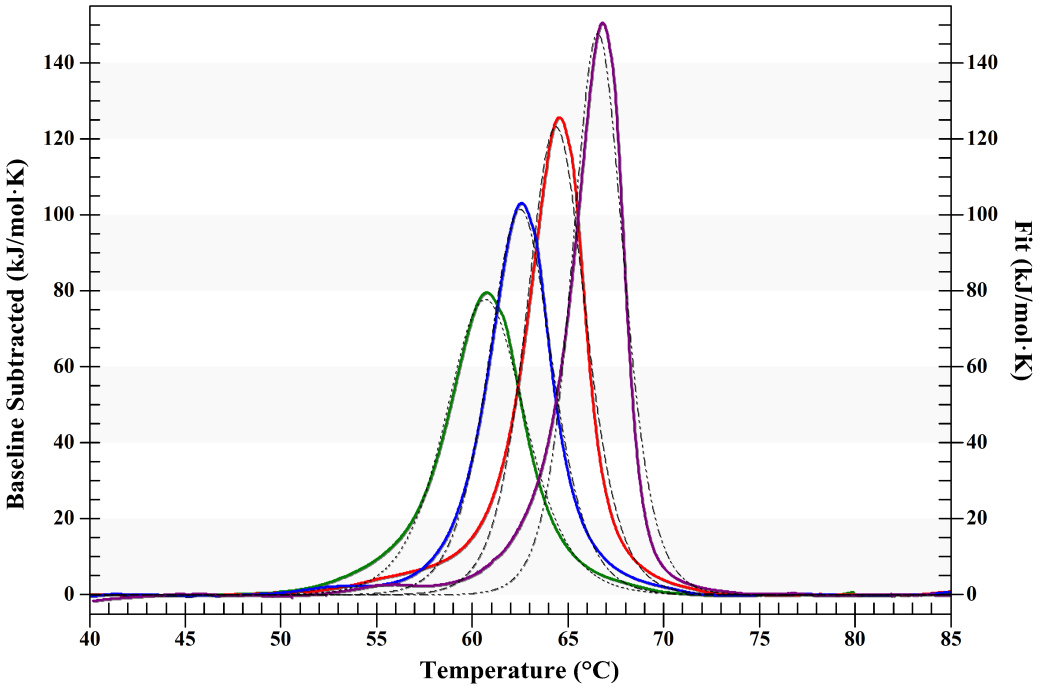


Figure 2.12: DSC analysis of *Cje*ATP-PRT. DSC measurements were performed under standard assay conditions with a $MgCl_2$ concentration of 5 mM. The resulting melting curves were baseline subtracted (solid lines) and fitted against a single melting event (dashed lines). Lines are coloured according to the ligand present in the experiment: none (green), histidine (blue), AMP (red), histidine and AMP (purple).

Table 2.6: Thermostability of *Cje*ATP-PRT. Melting temperatures were determined by DSF and DSC (sections 6.1.37 and 6.1.38) in the presence of different Mg^{2+} concentrations and known ligands. ND stands for not determined.

Additives	DSF	DSC
No ligand, no Mg^{2+}	54.4 ± 0.2 °C	ND
No ligand, 0.5 mM Mg^{2+}	56.7 ± 0.1 °C	ND
Mo ligand, 5 mM Mg^{2+}	57.5 ± 0.1 °C	60.7 ± 0.1 °C
2 mM ATP, 5 mM Mg^{2+}	57.8 ± 0.1 °C	ND
2 mM AMP, 5 mM Mg^{2+}	61.1 ± 0.3 °C	64.4 ± 0.2 °C
1 mM histidine, 5 mM Mg^{2+}	57.6 ± 0.1 °C	62.5 ± 0.1 °C
2 mM AMP, 1 mM histidine, 5 mM Mg^{2+}	62.3 ± 0.1 °C	66.6 ± 0.1 °C

2.8 Oligomeric state analysis

2.8.1 Analytical size exclusion chromatography

The oligomeric state of purified *Cje*ATP-PRT has been determined by analytical size exclusion chromatography (section 6.1.26) at 1.0 and 0.05 mg mL^{-1} enzyme concentrations at pH values of 8.5 and 7.5. The outcome was identical in all cases. A single peak eluted from the column at 11.4 mL corresponding to a mass of 204 kDa when subjected to linear regression obtained from the protein standards, which is in very close agreement with the theoretical mass of a hexamer (202 kDa).

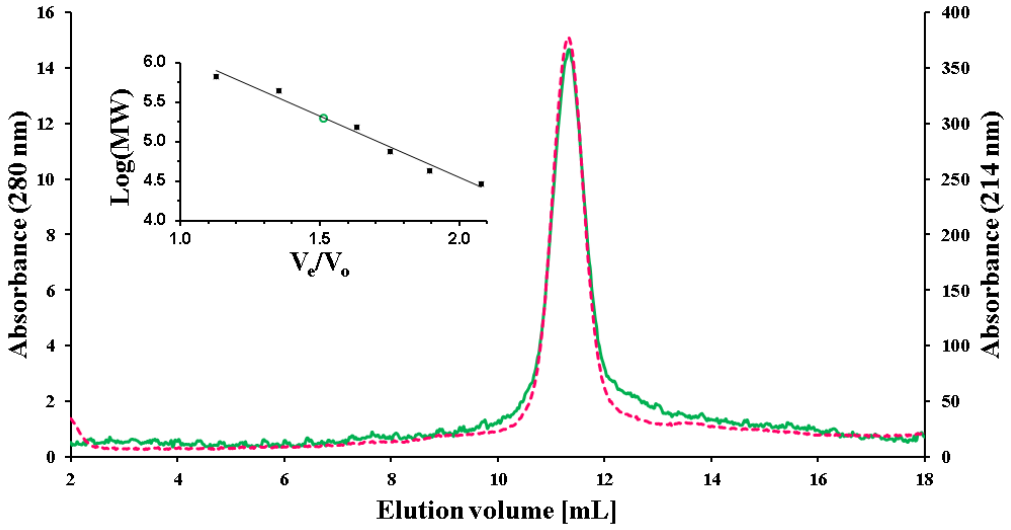


Figure 2.13: Low concentration analytical size exclusion chromatography run of *Cje*ATP-PRT. SEC was performed on 0.05 mg ml $^{-1}$ *Cje*ATP-PRT using a Superdex^{TM} 200 10/300 GL column as described in section 6.1.26. The absorbance of the protein sample was monitored at 280 nm (green, solid) and 214 nm (pink, dashed) due to the low extinction coefficient of *Cje*ATP-PRT at 280 nm. Inset: Calibration curve. The molecular weight of the standards proteins (Log(MW)) is plotted on a logarithmical scale against their normalised elution volumes (V_E/V_0). The void volume (V_0) was determined using Blue Dextran. The calculated molecular weight of the *Cje*ATP-PRT species is displayed as an open circle.

2.8.2 Static light scattering

A similar technique used to probe the quaternary structure of an enzyme is static light scattering (SLS). This method effectively combines size exclusion chromatography with the measurement of refractive index (RI), light scattering (LS), and viscosity (section 6.1.40). As observed in the analytical size exclusion experiment, only a single species was detected corresponding to a molecular mass of 188 kDa (Figure 2.14), which is slightly below the theoretical molecular mass of the hexamer, 202 kDa.

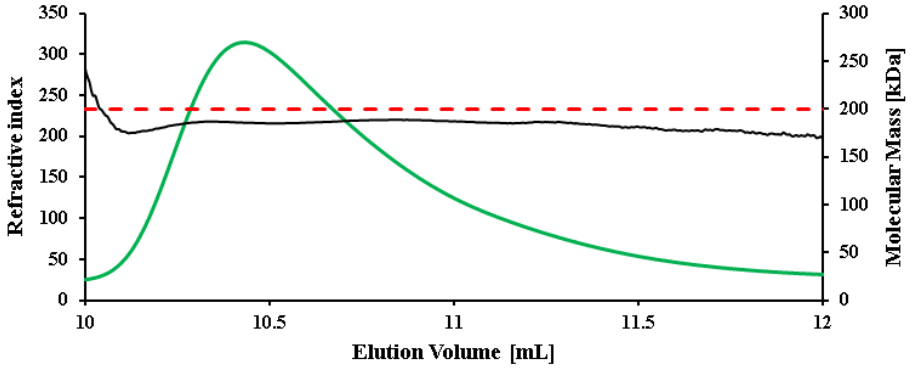


Figure 2.14: Static light scattering of *Cje*ATP-PRT. The refractive index (solid green line) of 1 mg mL $^{-1}$ *Cje*ATP-PRT was determined as described in section 6.1.40. The molecular weight of the mono-disperse species (black) eluting between 10 and 12 mL was calculated, using bovine serum albumine (BSA) as standard, and compared to the theoretical molecular mass of the *Cje*ATP-PRT hexamer (red dashed line).

2.8.3 Analytical ultracentrifugation

Due to the low extinction coefficient of *Cje*ATP-PRT around 280 nm it was necessary to perform all AUC experiments at 230 nm detection. An initial run was performed with concentrations ranging from 1.2–0.13 mg mL^{-1} which confirmed the results from SEC and SLS experiments. Only one species was detected with a sedimentation coefficient of 7.8, which corresponds to a molecular mass of 186 kDa.

The advantage of AUC over all other used methods is its high sensitivity. To decisively confirm the presence of the hexameric species even at very low protein concentrations another dataset was collected, at concentrations ranging from 0.1 to 0.025 mg mL^{-1}, with identical outcome. The unnormalised size distribution data (Figure 2.15) shows only one major species present in the sample. Using the SEDFIT parameters the obtained sedimentation coefficient of 7.9 was converted to a molecular mass of 182 kDa. The three minor peaks at 3.0 S, 4.8 S and 11.2 S were converted to molecular masses 40 kDa, 87 kDa and 320 kDa accordingly, which would be closest to the masses of a theoretical monomer (33.7 kDa), dimer (67.4 kDa) and de-

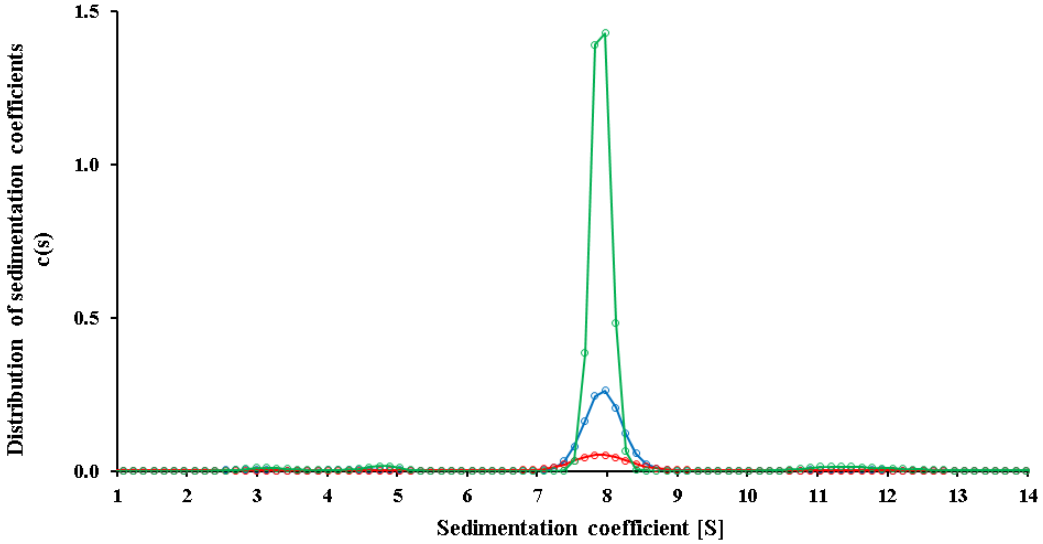


Figure 2.15: Size distribution function of *Cje*ATP-PRT. AUC experiments were conducted according to section 6.1.41. The resulting data was analysed using SEDFIT.[90] The plot shows the data of all three used enzyme concentrations unnormalised from highest to lowest signal: 0.1 mg mL $^{-1}$ (green), 0.05 mg mL $^{-1}$ (blue) and 0.025 mg mL $^{-1}$ (red).

camer (337 kDa), but did not change in relation to the main peak. Since the lowest measured concentrations is about two times higher than that used in all kinetic assays, it is evident that the vast majority of *Cje*ATP-PRT exists as a hexamer in solution under the given conditions.

2.9 Crystal structure

2.9.1 Crystallisation

During the initial crystallisation trials (section 6.1.44) *Cje*ATP-PRT did not crystallise at high protein concentration due to the formation of heavy precipitate. Instead, the optimal protein concentration was found to be 2–4 mg mL^{-1} where small or thin crystals would form in multiple conditions of both PACT and JCSG^{+} screens. Unfortunately none of the conditions yielded crystals suitable for diffraction. Histidine was therefore added to the

crystallisation conditions at a concentration of 1 mM, which resulted in the formation of cuboid-shaped crystal in PACT screen condition F4 (section 6.2.5). Fine-tuning of this condition delivered the following crystallisation condition: 100 mM BTP pH 6.5, 200 mM KSCN, 1 mM histidine and 13–14 % w/v polyethylene glycol (PEG) 3350. Crystals formed overnight and continued to grow for two days after the drops were put down (Figure 2.16). The chosen cryoprotection procedure preserved the crystal appearance. All diffraction data was acquired at the MX beamlines of the Australian Synchrotron using the cryojet and a constant beam energy of 13000 eV.

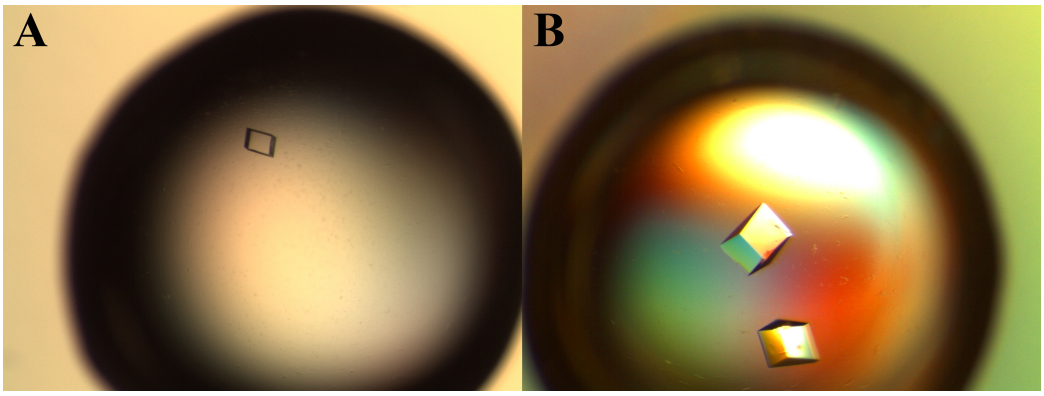


Figure 2.16: *Cje*ATP-PRT crystals. Photographs taken of drops containing *Cje*ATP-PRT crystals. A: histidine cocrystal, condition derived from PACT F4: 0.1 M BTP, 0.2 M KSCN, 13–14 % w/v PEG 3350, 1 mM histidine, pH 6.5. B: ATP cocrystals, condition derived from Clear Strategy I A9: 0.1 M Na acetate, 0.1 M $MgCl_2$, 13–15 % w/v PEG 4000, 10 mM ATP, pH 5.5.

After the initial success the screening for suitable crystallisation conditions continued focussing on co-crystallisation with AMP and ATP as well as a ligand-free form. Crystals were grown successfully from several additional conditions (section 6.2.5). Diffraction data was acquired for two of these, featuring ligand density for ATP, and AMP and histidine (Table 2.7). Promising diffraction data was also collected from a ligand-free crystal, but unfortunately due to the low crystal symmetry and an error during collection.

2.9.2 Molecular replacement strategy

The unit cell and symmetry ($P12_11$) obtained were different from the previously published ATP-PRT structures of *E. coli*, *M. tuberculosis,* and *M. thermautotrophicus*. As a result early attempts to solve the structure by molecular replacement were unsuccessful. All attempts made to obtain a heavy atom data set to solve the phases also failed. Neither the soaking in bromide nor the use of a xenon pressure chamber resulted in suitable diffraction from formed crystals and the substitution of thiocyanate in the crystallisation condition with selenocyanate failed to generate crystals.

The phases were finally solved by applying a complex molecular replacement strategy. It was expected from the unit cell dimensions of the initial (Histidine bound) dataset that six complete chains were present. Initially only domain I and II of the *E. coli* ATP-PRT structure (PDB code: 1H3D) were used to find a suitable dimer. This dimer was then used as model in a consecutive search resulting in a double circular arrangement of the six truncated chains that largely resembled the hexameric structure published previously.[26] Starting from this solution, the ACT domains were manually built into the density on all six chains through multiple refinement rounds resulting in the first complete hexameric structure of *Cje*ATP-PRT (PDB code: 4YB5). While the completely refined hexamer was successfully used as a search model for molecular replacement to solve the AMP and histidine bound structure (PDB code: 4YB6), it was found unsuitable for the dataset containing ATP in spacegroup $P1$. By applying the same strategy as above, in this case 12 complete chains, ergo two whole hexamers (overall RMSD = 0.563 Å), were resolved (PDB code: 4YB7). Three effector molecules were observed in the crystal structures of *Cje*ATP-PRT. While both AMP and histidine bound ATP-PRT long form structures have been described previously, the binding of ATP is unprecedented in the ATP-PRT literature. In all cases, the ligand molecules were added last into the density maps during refinement to avoid the introduction of model bias.

Table 2.7: Crystal parameters, data collection, and refinement statistics.

	Histidine	Histidine/AMP	ATP
		Data collection	
Crystal system	monoclinic	monoclinic	triclinic
Space group	$P12_11$	$P12_11$	$P1$
Unit cell parameters			
a, b, c [Å]	91.14, 123.22, 95.70	91.87, 124.91, 92.81	91.67, 91.83, 154.90
α, β, γ [°]	90.00, 110.66, 90.00	90.00, 115.86, 90.00	101.11, 95.21, 118.14
Resolution range [Å]	89.55–2.24(2.28–2.24)	50.00–1.98(2.01–1.98)	50.00–2.20(2.24–2.20)
Measurements	472477	493057	839825
Unique reflections	94235	127994	212958
Completeness	89.9 (89.7)	98.1 (96.9)	91.8 (97.3)
Redundancy	5.0 (4.5)	3.9 (3.9)	3.9 (4.0)
$I/\sigma(I)$	9.9 (2.4)	14.6 (2.9)	9.8 (1.8)
R_{merge}	0.093 (0.795)	0.059 (0.422)	0.076 (0.841)
$CC_{1/2}$	0.997 (0.625)	0.998 (0.837)	0.982 (0.568)
Wilson B value [Å^2]	43.8	21.3	36.8
Matthews coefficient	2.48	2.37	2.72
		Refinement	
Resolution [Å]	48.04–2.24	41.34–1.98	48.06–2.20
R_{cryst}	0.211	0.203	0.236
R_{free}	0.241	0.220	0.259
Chain length	300	300	300
Observed residues	1753	1758	3531
Water molecules	128	714	365
Others	6 histidine, 8 SCN^-, 5 K^+, 3 PEG	6 histidine, 6 AMP, 9 Mg^{2+}, 6 PEG	12 ATP, 12 Mg^{2+}, 6 PO_4^{3-}, 1 acetate
Mean B [Å^2]			
Protein	50.03	24.58	47.13
Water	38.62	26.94	35.20
Other	61.30	45.09	55.83
Ligand	34.64	13.06, 24.81	51.27
RMSD			
Bond lengths [Å]	0.012	0.009	0.009
Bond angles [°]	1.465	1.372	1.470
Chiral volumes [Å^3]	0.075	0.077	0.073
Ramachandran			
Preferred %	98.56	98.36	98.22
Allowed %	1.44	1.64	1.78
Outliers %	0.00	0.00	0.00
PDB entry	4YB5	4YB6	4YB7

2.9.3 General fold

The crystal structures of *Cje*ATP-PRT contain six or 12 identical protein chains in the unit cell. Each chain shows the same fold as previously determined structures, composed of 10 α helices, 16 β strands and the connecting loops arranged in three discrete domains: domain I, II, and III.[26,35] A schematic of the *Cje*ATP-PRT topology is shown in Figure 2.17.

Domain I is composed of residues 0–103 and 191–225, which form an α-β open twisted sheet structure, or Rossmann fold, commonly found for nucleotide binding proteins.[91] In detail it consists of a central twisted six-stranded β sheet with four parallel (β1, β3, β4 and β5) and two anti-parallel (β2 and β11) β strands, as well as four flanking α helices, two on one side of the β sheet (α1 and α8) and two on the other (α2 and α3). The first few N-terminal residues are unresolved in all three *Cje*ATP-PRT structures and the β5 strand appeared in some chains as a loop running in close proximity along the β11 strand.

A similar fold is adopted by domain II, comprising residues 104–190, but the central five-stranded β sheet (strands β7, β8, β9 and β10 parallel, and β6 anti-parallel) are strongly curved and curled around α4 and α5 on one side while α6 and α7 sit almost parallel next to each other on the other side. Domain II is inserted into domain I at the primary sequence level into the loop between β5 and β11 and consequently the β5-β6 loop and the β10-β11 loop form the only two covalent connections of the two domains. These loops also form one boundary of the active site.

Domain III is formed from the 74 C-terminal residues (225–299) and possesses the distinctive features of an ACT domain, an anti-parallel four-stranded β sheet (β12 and β13, β14 and β15) with two α helices (α9 and α10) on one side. The special feature of the ATP-PRT ACT domain is the last C-terminal β strand β16, which sits almost perfectly orthogonal to β15 and is donated into the ACT domain of the adjacent chain, effectively constructing a five-stranded β sheet. Each ACT domain also houses the two faces of the allosteric binding site for the feed-back inhibitor histidine, α9-β13

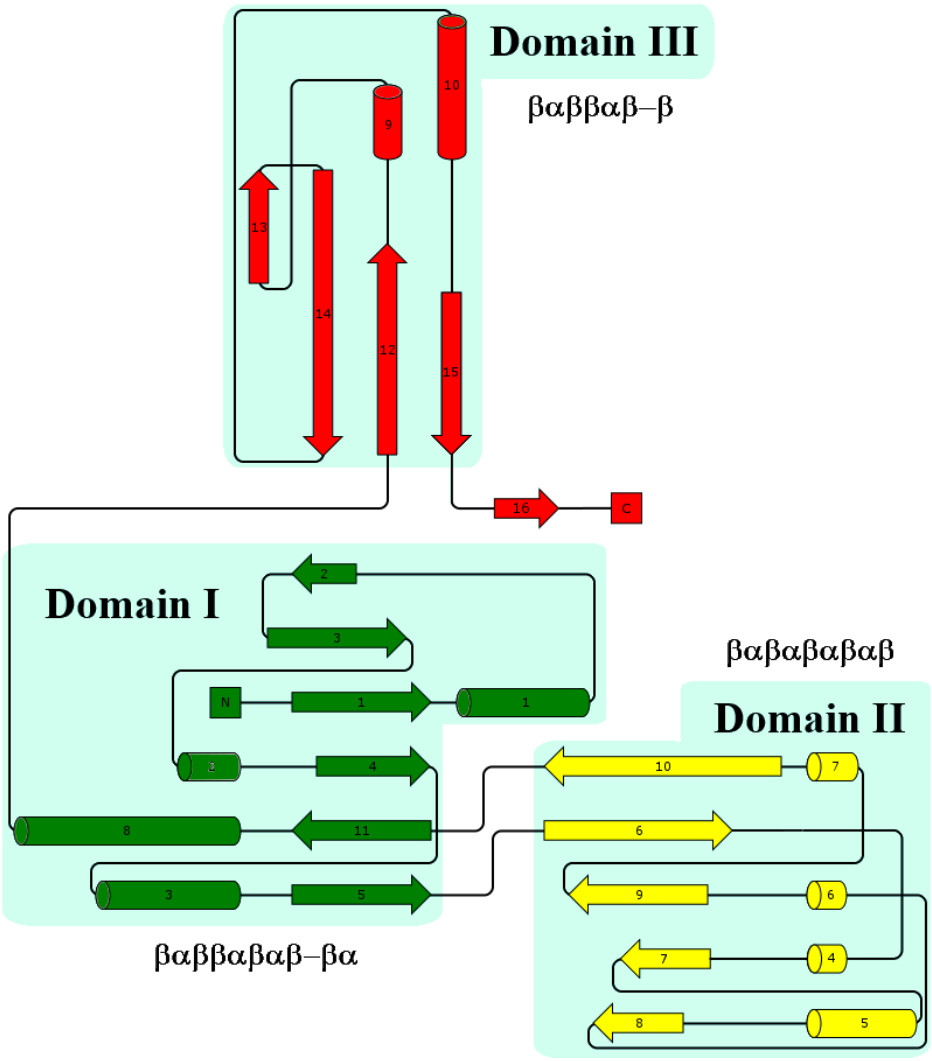


Figure 2.17: Secondary structure of the *Cje*ATP-PRT single chain. The three domains of *Cje*ATP-PRT as 2D cartoon representation. The size of α helices and β strands is proportional to the number of incorporated residues while loop/connection length is random. Secondary structure elements are coloured according to domain association: domain I (green), domain II (yellow), domain III (red). Image was created using Pro-origami.[92]

loop on one and $\beta 12/\beta 15$ on the other side. Its inter-molecular interaction and allosteric binding site clearly define the *Cje*ATP-PRT ACT domain as a functional regulatory domain.[41]

The *Cje*ATP-PRT single chain is grossly shaped like a triangular wedge with domains I and II positioned close together and the ACT domain sitting aside on an angle relative to the domain I - domain II axis. The ACT domain is connected to the catalytic domain I by the long α8 helix and the short α8-β12 loop, allowing for a certain degree of flexibility. The large active site is situated in the cleft between domains I and II and recruits residues from both domains of one chain. The active site ligands found in the structures 4YB6 and 4YB7 revealed some detail about the *Cje*ATP-PRT active site architecture. The active site can be largely split into two regions: the ATP-binding groove of domain I and the PRPP binding site of domain II (Figure 2.18), with each area sitting on the edge of the beta sheet of the corresponding domain. The ATP-binding site the ATP bound *Cje*ATP-PRT includes the coordination site for a single Mg^{2+} ion between two aspartate side chains (D55 and D56) and the β and γ phosphate of ATP.

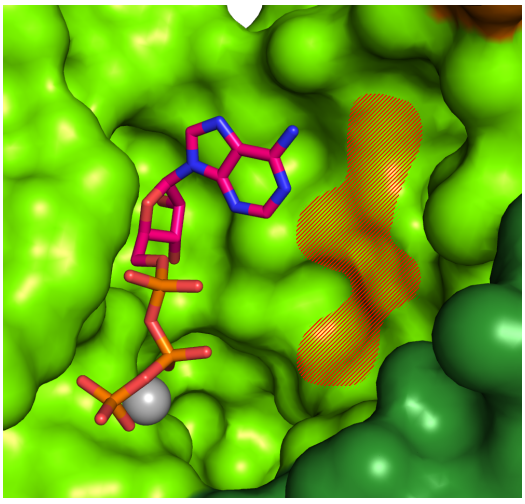


Figure 2.18: The *Cje*ATP-PRT active site. The active site of a *Cje*ATP-PRT (4YB7) chain (light green), adjacent to its dimer partner chain (dark green) and the α1-β2 loop of a neighbouring chain (brown), is represented as a surface. Domain I is to the left of the binding site and domain II to the right. Residue R16 has intentionally been removed for better visibility as it is positioned directly above the adenine. ATP (pink) is displayed as sticks with the coordinated Mg^{2+} ion (grey sphere) in the active site. The expected binding site of PRPP is represented by red hatching. Hetero-atoms of ATP are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

2.9.4 ATP bound structure

The structure of the ATP bound *Cje*ATP-PRT (4YB7) contains 12 individual chains in the unit cell, forming two homo-hexameric assemblies (Figure 2.19). Two chains each form a dimer with an approximately 180° rotational symmetry. The interdomain contacts are made between the catalytic domains and are located on the opposite side to their intramolecular domain connection (β5-β6 and β10-β11 loops). This effectively combines the two active sites of the dimer into one large cavity. The dimer interface buries an area of approximately 1000 Å^2 in all six observed dimers, corresponding to 7 % of the overall surface of a single chain. The main contacts at the dimer interface are made between α6 and its flanking loops (domain II) of one chain with β1 and the β3-α2 loop (domain I) of the other chain. The interactions of α7 (domain II) with the α1-β2 loop (domain I) also make a small contribution. Due to the rotational symmetry, these interactions are inversely repeated on the other side of the interface.

Three dimers in turn come together to form the hexamer. The dimers attach side by side along their ACT-to-ACT axes, with their active sites on the inside of the formed complex. The hollow-centred core structure is held together by two trimeric ACT domain units, one on each end. Two neighbouring chains in such a unit bury approximately 850 Å^2 of surface area between them (6 % of their total surface), with the predominant contribution made by the ACT-ACT interface. The main points of interaction here are the tight contacts between the C-terminal β16 and β15 and α10 of the neighbouring chain, as well as contacts between α8 of one and α10 of the other chain. According to the analysis of the structure with PDBe PISA[93] this second interface, referred to as trimer interface, was classified as essential for the complex with a higher complex formation significance score (CSS) than the dimer interface (Table 2.8). This is a strong indicator for the presence of the hexamer as permanent structural and functional entity.

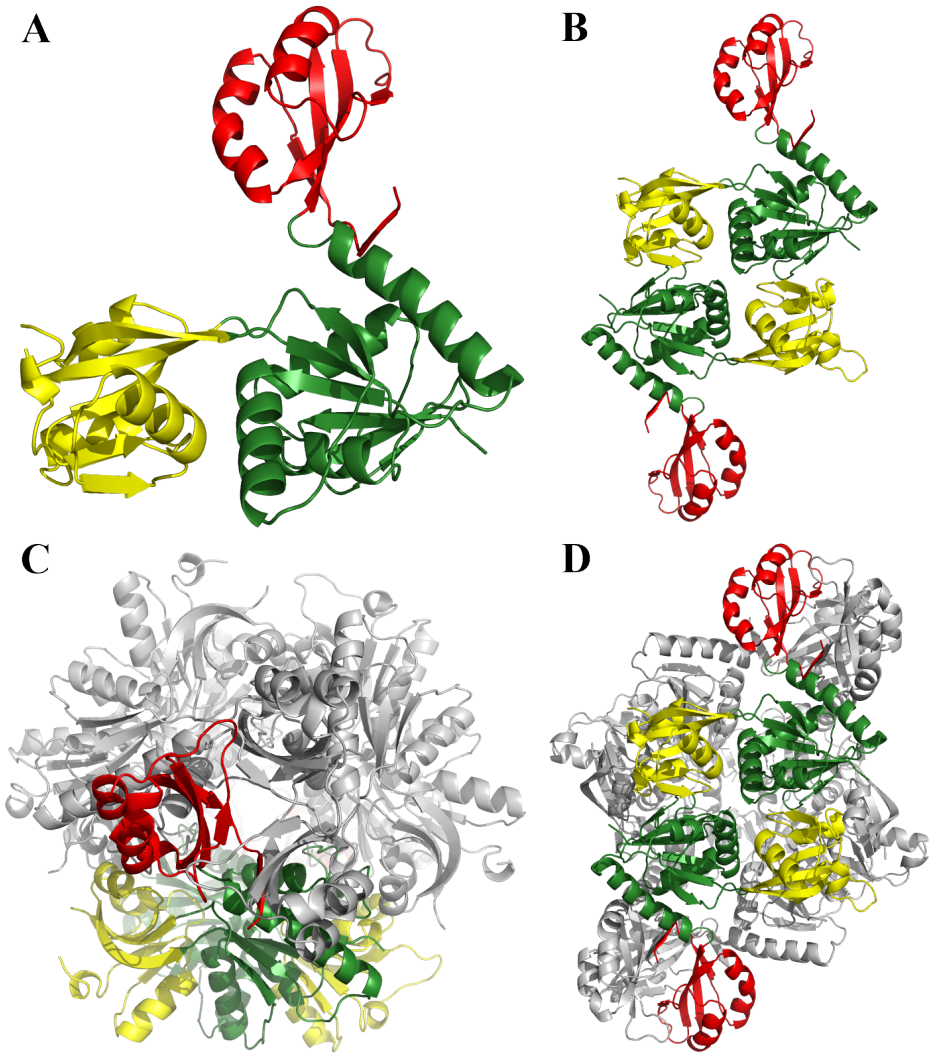


Figure 2.19: Structure and multimeric assembly of *Cje*ATP-PRT. Crystal structure of the histidine bound *Cje*ATP-PRT complex (4YB5) in the cartoon representation as example for the general multimeric assembly of *Cje*ATP-PRT. A single chain is structurally divided into domain I (green), domain II (yellow) and domain III or ACT domain (red). A: Monomer B: Hypothetical dimer C: Hexamer, view along the trimer axis D: Hexamer, view along the dimer axis.

There are six main triangular-shaped openings to the central cavity of the hexamer, each between the β5-β6 and β14-α10 loops of one chain and the α1-β2 loop of the neighbouring chain. The centre of each dimer interface also contains a small opening.

The binding of the substrate molecule ATP in structure 4YB7 (Figure 2.20) was observed in the nucleotide binding groove of domain I, involving a complicated network of hydrogen bonding interactions with a small number of residues that were in part found to be conserved (Figure 2.3). The most obvious interactions can be seen for the binding of the triphosphate moiety of ATP, which is flanked by two arginine side chains. R54 interacts with the γ phosphate while R16 forms hydrogen bonds to the oxygens of the α phosphate. Q12 also reaches into the active site, which places its side chain nitrogen in an optimal position to interact with several of the triphosphate oxygens. Interestingly Q12 is only conserved within the enterobacterial ATP-PRTs. All other characterised enzymes possess a proline residue in this position, making this interaction unlikely.

Another striking feature is the presence of a Mg^{2+} in an octahedral coordination. It is coordinated by the side chains of the successive aspartate residues D55 and D56 at the start of α2, two oxygen atoms, one each from the β and γ phosphate, and two water molecules. This coordination both embeds the Mg^{2+} into the binding site of *Cje*ATP-PRT and provides strong interaction with the ATP triphosphate.

N75 is the only residue found within hydrogen bonding distance to the ATP ribose. The side chain of this functionally conserved residue forms hydrogen bonding interactions to both ribose oxygens simultaneously. The ribose ring itself is inserted into a small cavity created by the backbone and short sidechains of L17, G73, G102, S191, R192 and A193. Despite the relative hydrophobicity of this pocket, water molecules were observed surrounding the ribose 2' and 3' oxygens.

The adenine portion of ATP forms a single hydrogen bond to S172, which is part of the highly conserved PRPP binding loop. It is further held in place by hydrophobic interactions with the side chain of R16, the binding of which to the α phosphate has been described, on one side and L170 and C104 on the other side. This "molecular clamp" is not as rigid as the rest of the binding interactions, allowing the adenine to adopt two different orientations in the active site (Figure 2.21). In each of the 12 active

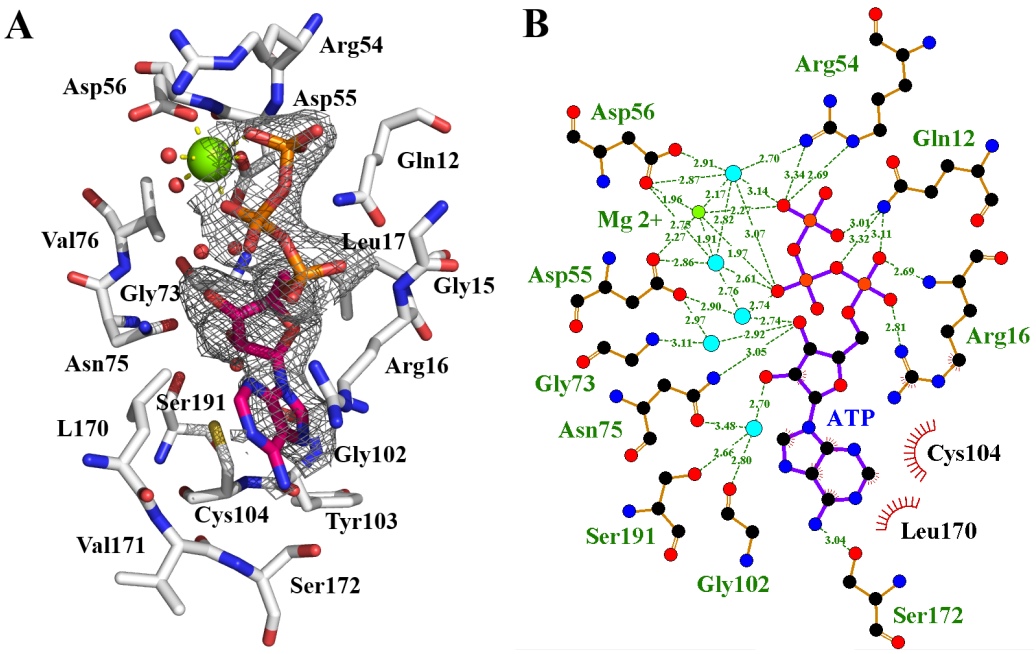


Figure 2.20: ATP binding site of *Cje*ATP-PRT. Observed binding mode of the ATP molecule found in 4YB7. A: Real space representation of the ATP molecule (pink) surrounded by the observed electron density (Fo-Fc map - grey mesh) and the active site residues (white) as sticks. B: Planar plot of all ATP interactions created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with distances. Hydrophobic interactions are displayed as short radial red lines. Water molecules are displayed as cyan spots. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), magnesium (green).

sites the ATP-adenine was observed in either a catalytical relevant position, placing the N_1 nitrogen close to the expected PRPP binding site, or flipped by a 180°rotation. Overall the observed occupancy for ATP is comparatively poor for the adenine ring portion in all 12 chains, reflecting its rotational freedom. Although it appears very prominent, the importance of residue R16 for ATP binding is undermined by the fact that it is only partially conserved, implying an auxiliary role only.

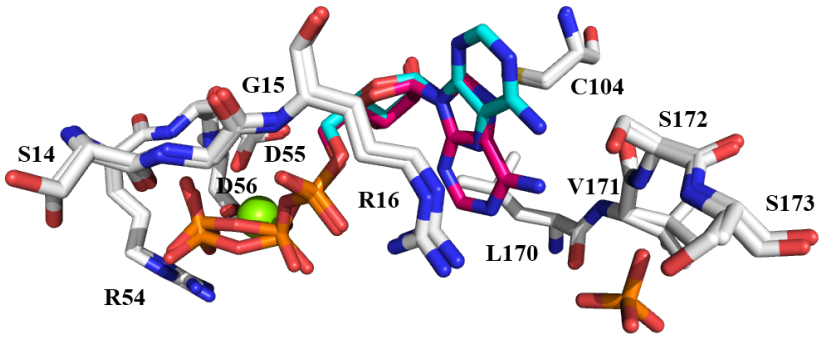


Figure 2.21: Alternative ATP binding modes in *Cje*ATP-PRT. Superposition of the two observed ATP binding modes found in 4YB7 surrounded by the active site residues (white), in stick representation. The catalytically relevant orientation (pink) differs from the catalytically irrelevant position (cyan) by a flip of the adenine portion of the molecule. A phosphate is found in chains displaying the irrelevant ATP binding mode. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), magnesium (green).

2.9.5 Histidine bound structure

The histidine bound crystal structure of *Cje*ATP-PRT contains six individual chains in the unit cell forming a single hexameric assembly. Overall, the 2-fold and 3-fold symmetry, as seen in each hexamer of 4YB7, is retained with all active sites facing the inside of the holo-enzyme complex (Figure 2.19).

The dimer interface is, as in the ATP bound crystal structure, created by cross-over contacts between the catalytic domains of the two involved chains, but buries a larger surface area, 1200 Å^2, approximately 8 % of the total surface of a single chain. The main contacts involve residues of α6 and its flanking loops, as well as β7 and β8 on one chain, and residues in β1, β2, β3 and α2 on the other chain. Helix α7 is not involved in the dimer interface in 4YB5 at all, but the C-terminal ends of α3 of both chains come together in the centre of the interface. The trimer interface in 4YB5 buries about the same area as the dimer interface (1200 Å^2), underlining the importance for both interfaces in the holo-enzyme. Each interface also includes a single histidine and a conserved water molecule, which lend additional contacts to the overall tighter interface.

The six triangular openings into the central cavity found in structure 4YB7 are also present in 4YB5 and are only marginally smaller (10 ±1 vs. 12 ± 1 Å in width). Due to rearrangements of the dimer interface (as described in greater detail in section 2.9.7), the small opening between the two dimer chains is completely closed.

Histidine binding in structure 4YB5 occurs at the interface between two adjacent ACT domains in their trimeric arrangements. This means that one histidine is bound per chain but each chain provides two different faces for interaction with histidine. The histidine carboxy and amino groups are nestled into a highly conserved binding loop with the consensus sequence PGXXXPT between α9 and β13 of one ACT domain, while the imidazole side chain is inserted into a small cavity created by the side chains on the β sheet of the other ACT domain. This shared binding site allows the bound histidine molecules to act like pins that pull the ACT domains together, creating the conformational changes described later.

Looking at the histidine binding loop of *Cje*ATP-PRT in closer detail reveals the coordination of the histidine carboxy oxygen atoms by four backbone nitrogen atoms, three from the residues V248, E249 and R250 in the conserved binding loop, and one contributed by residue V268 on the nearby passing β strand β14. The amino group of histidine forms interactions with the backbone oxygen of P251 and the side chain oxygen of T252 (Figure 2.22). This absolutely conserved threonine residue has been shown to be crucial for histidine binding,[77] but it is the two proline residues P246 and P251 paired with the flexibility of G247, which provide this little loop with the narrow twist that is required for the observed conformation.

The binding cavity on the other side is formed by residues M230, H232, S288 and L290. It can be roughly split in two short motifs. The first is situated on β12 and accommodating the residues M230 and H232 (AXB motif) and the second is part of β16 and made up by residues S288 and L290 (CXL motif). Both motifs are structurally well conserved, as the overall fold of the ACT domain β sheet is consistent in all deposited structures, but the sequence allows a certain variability of the participating residues. In the AXB

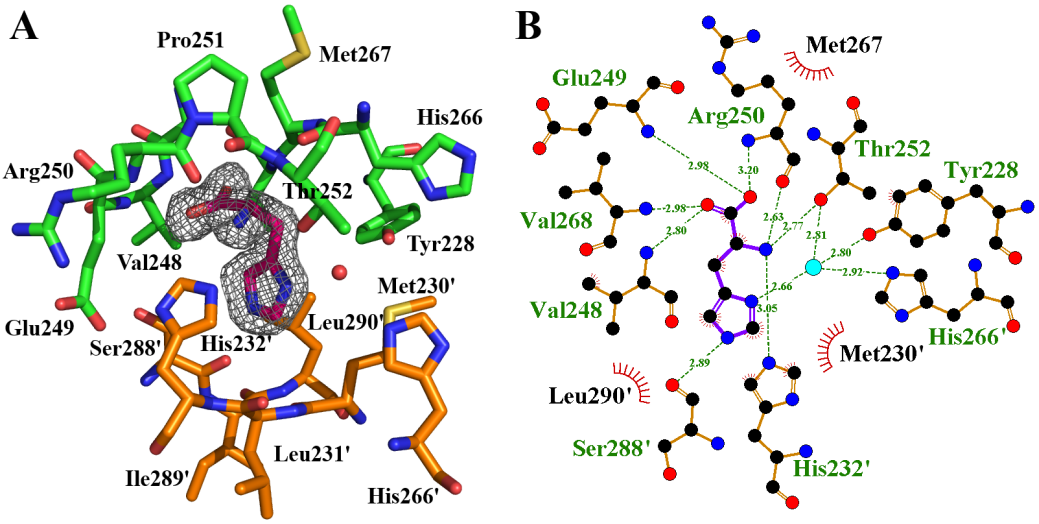


Figure 2.22: Histidine binding site of *Cje*ATP-PRT. Histidine binding mode found in 4YB5. A: Real space representation of the histidine molecule (pink) surrounded by the observed electron density (Fo-Fc map - grey mesh) and the residues of adjacent ACT domains (green and orange) as sticks. B: Planar plot of all histidine interactions created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with distances. Hydrophobic interactions are displayed as short radial red lines. Water molecules are displayed as cyan spots. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), sulphur (yellow).

motif, positions A and B can be populated by valine, methionine, aspartate, asparagine or histidine. The CXL motif on the other hand appears to be more consistent throughout species with C always being a residue with a short side chain (either glycine, alanine, serine or aspartate), and the leucine being absolutely conserved.

Despite the changing residue environment, the binding mode of histidine in *Cje*ATP-PRT is identical to the binding mode presented in *Mtu*ATP-PRT (1NH8) and *Mth*ATP-PRT (2VD3). Therefore a medium side chain length and limited polarity appear sufficient to bind the imidazole side chain of histidine, which is found stacking with the side chain of either aspartate, asparagine or histidine, but specific functionalities are not required. The importance of the mostly sterically conserved motifs AXB and CXL is supported further by the finding that the substitution of residue N215 and A270

in *Cgl*ATP-PRT (equivalent to H232 = "B" and S288 = "C" in *Cje*ATP-PRT) to lysine and proline respectively resulted in a lower sensitivity to histidine, with 50 % and 35 % residual activity in the presence of 2 mM histidine compared to 10% of the *Cgl*ATP-PRT wild type.[77]

2.9.6 AMP and histidine bound structure

While equivalent to 4YB5 in conformation, arrangement and histidine binding, the structure 4YB6 features the inhibitor AMP, which was found to interact almost entirely with residues of the PRT signature fold in specific residues 169–176 (β9-α7 loop). Most residues of this highly conserved sequence surround the phosphate moiety in a half circle, presenting their backbone nitrogen atoms towards it. The short polar side chains of S172, S173 and T176 provide additional anchor points for the negatively charged phosphate (Figure 2.23). This arrangement allows for multiple (6–12) hydrogen bonding interactions to be formed, which give this binding site its high coordination strength. It is therefore unsurprising that it was found occupied by inorganic phosphate in the ATP bound structure (4YB7).

The most interesting finding is the coordination of H33, and potentially also H35, of the neighbouring subunit to the bound phosphate group of AMP. Histidine H33, along with H35, is positioned on the flexible α1-β2 loop that protrudes into the active site upon histidine binding and effectively "caps" the phosphate binding site.

The AMP ribose 2' and 3' oxygens are in close proximity to the conserved residues D169 and E156, with which they are likely to form hydrogen bonding interactions. As described for ATP binding, the adenine ring system is weakly coordinated by a hydrophobic stack between R16 and L170. The only notable difference is the inverted orientation of the purine ring system, which is a logical consequence of the 5'-phosphoribosyl (of AMP) and 5'-triphosphoribosyl (of ATP) groups being located on opposite sides. The side chain of R16 was unresolved in the experimental electron density likely due to the absence of the negatively charged triphosphate to hold it in place.

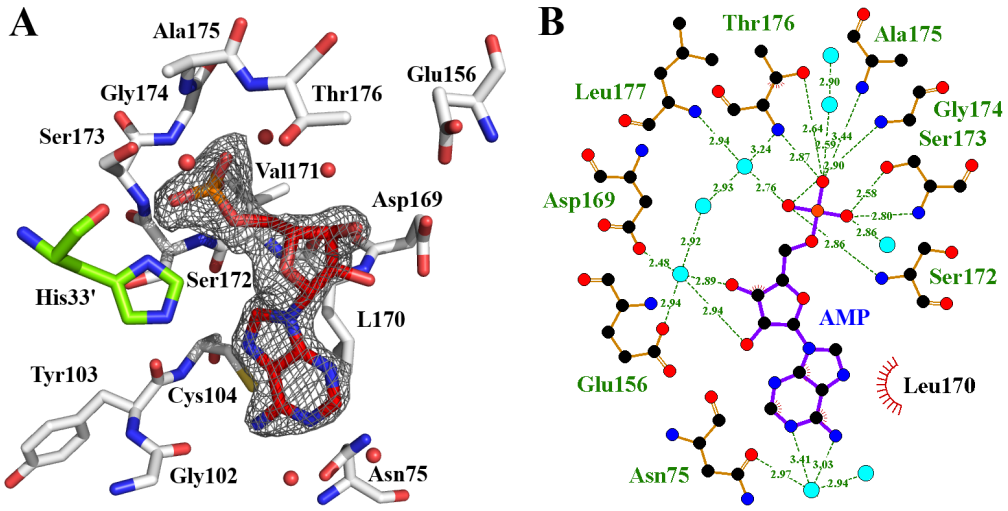


Figure 2.23: AMP binding site of *Cje*ATP-PRT. Observed AMP binding mode of the ATP molecule found in 4YB6. A: Real space representation of the AMP molecule (pink) surrounded by the observed electron density (Fo-Fc map - grey mesh) and the active site residues (white) plus H33' of the adjacent subunit (green) as sticks. B: Planar plot of all AMP interactions created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with the real space distances between atoms. Hydrophobic interactions are displayed as short radial red lines. Water molecules are displayed as cyan spots. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), sulphur (yellow).

The crystal structure 4YB6 reveals AMP binding occurs in both, the ATP and the PRPP binding region, of the active site. The 5'-phosphate and ribose are closely coordinated into the PRPP binding motif, with the adenine bridging across into the ATP binding site. The identical binding mode was observed in the crystal structure of *Eco*ATP-PRT (PDB code: 1H3D),[26] the only difference being a closer coordination by E156 and R16 and the absence of the cap (side chains of H33 and H35). In *Mtu*ATP-PRT (PDB code: 1NH8)[35] the bound AMP is flipped by 180°. The adenine and phosphate groups are found in approximately the same position as for *Cje*ATP-PRT, but the ribose is turned around and out of the active site. There it is making hydrogen bonding contact with the side chain of aspartate D30 of the neighbouring chain, which is equivalent to H35 in *Cje*ATP-PRT.

2.9.7 Comparison of active and inhibited conformations

*Cje*ATP-PRT exists crystallographically in two distinct forms as has been described for *Mtu*ATP-PRT previously.[35] The open, or active, conformation is observed in the absence of histidine. In the case of *Cje*ATP-PRT this conformation is represented by the ATP bound crystal structure (4YB7). The closed, or inhibited, conformation is found exclusively in the presence of the allosteric inhibitor histidine, bound at the trimer interface. Both remaining structures of *Cje*ATP-PRT (4YB5 and 4YB6) show the inhibited conformation (overall RMSD = 1.39 Å). The main difference between them is the presence/absence of AMP in the active site. Although it is a potent inhibitor, the binding of AMP to the active site alone does not induce the same conformational changes as histidine binding, as shown in *Eco*ATP-PRT. Instead the enzyme stays in the open form, ready to bind histidine.[26]

The differences between conformations can be seen in both the individual chain and the holo-enzyme. Ultimately, all conformational changes are facilitated by a rotational twist of the ACT domain connection by approximately 30° around the flexible α8-β12-hinge, in each chain (Figure 2.24A). As it is performed simultaneously six times per hexamer, this small movement restructures the entire complex.

Although the changes in a single chain are small (RMSD = 1.65 Å, between single chains of 4YB6 and 4YB7) and appear to affect the catalytic domains to an even lesser extent (RMSD = 1.22 Å, between residues 5–215 of 4YB6 and 4YB7 single chains), they significantly change the appearance of the holo-enzyme. Structural comparison of the *Cje*ATP-PRT hexamer in open and closed conformation clearly shows large discrepancies with average RMSD values of 35–40 Å between 4YB7 and 4YB5 or 4YB6, respectively. The open hexamer of *Cje*ATP-PRT (4YB7) is wider in the central catalytic core (96 ± 1 Å vs. 92 ± 1 Å) and shorter when measured from one ACT-trimer to the other (117 ± 1 Å vs. 122 ± 1 Å), compared to the closed conformation (Figure 2.24B). The observed overall changes can be retraced

to changes in the two interfaces presented in *Cje*ATP-PRT hexamers. The analysis of these interfaces of all three crystal structures revealed significant differences (Table 2.8).

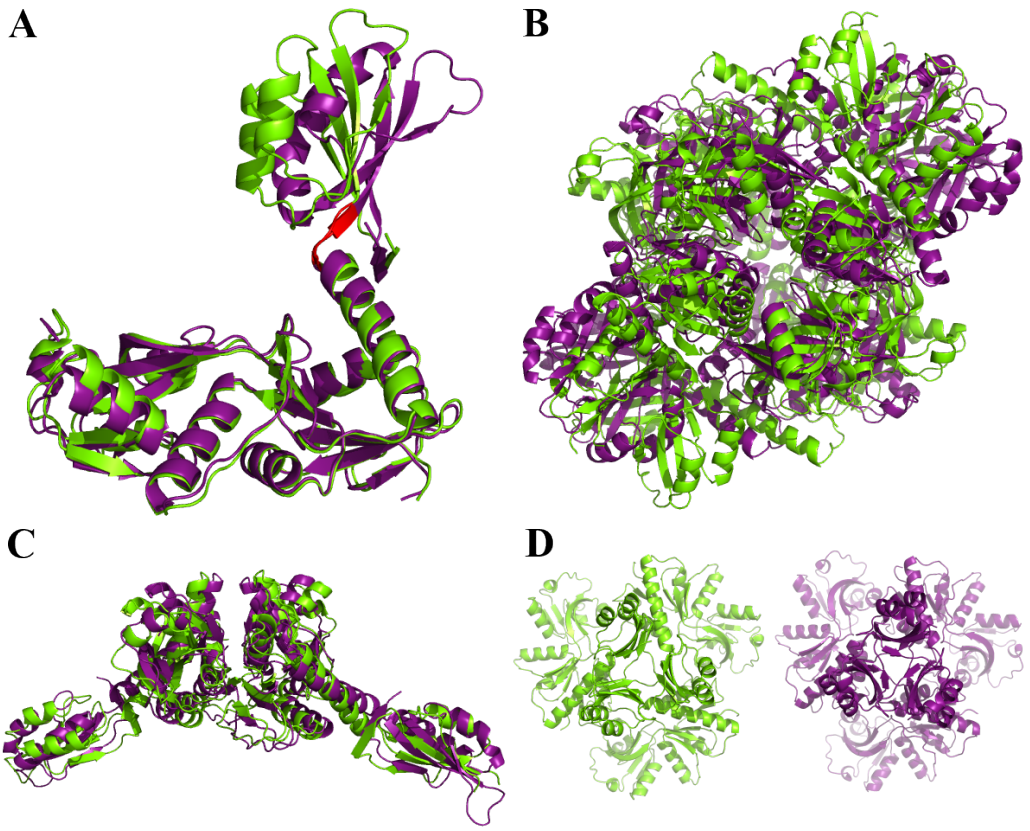


Figure 2.24: Conformational changes caused by histidine binding. A series of superimpositions of the open (4YB7) (green) and closed (4YB5) (purple) conformations of *Cje*ATP-PRT are shown in the cartoon representation to highlight the main differences between them. A: Superimposition of the single chain on the catalytic core domains to visualise the relative ACT domain movement. The α8-β12-hinge is highlighted in red. B: Superimposition of the complete hexamers shows large discrepancies in chain arrangement. C: Superimposition of a dimeric arrangement shows a "flex" of the dimer interface. D: The open and closed ACT domain arrangements side by side. The trimeric structure rotates and compacts in the presence of histidine.

Table 2.8: *Cje*ATP-PRT interfaces. Interfaces presented in the three crystal structures of *Cje*ATP-PRT were analysed using PDBe PISA.[93] Values are given as average of all individual interfaces.

Structure	Interface	Buried area [Å^2]	[%] of total area	CSS
4YB5	Dimer	1181	7.8	0.361
4YB6	Dimer	948	6.3	0.327
4YB7	Dimer	996	6.5	0.377
4YB5	Trimer	1220	8.0	0.598
4YB6	Trimer	1223	8.1	0.619
4YB7	Trimer	827	5.5	0.538

The trimer interface, which is directly influenced by the binding of histidine, is significantly larger in 4YB5 and 4YB6 with an average buried surface of 1200 Å^2 compared to 830 Å^2 in 4YB7. Histidine is positioned in the centre of the interface interacting closely with residues V248–T252 on one chain and H232 on the other, bringing those residues closer together. This can be quantified by the distances between the α carbon atoms of E249 and H232' (8.8 $\pm$ 0.1 Å vs. 11.6 $\pm$ 0.3 Å), and T252 and A264' (6.9 $\pm$ 0.1 Å vs. 10.6 $\pm$ 0.3 Å) in 4YB6 and 4YB7, respectively.

Additionally, a conserved water molecule was observed in all histidine binding sites in both inhibited *Cje*ATP-PRT structures (4YB5 and 4YB6). It is positioned strictly beside the δ1 nitrogen atom of the bound histidine molecule, able to form additional hydrogen bonding interactions with T252 and Y228 of one chain and H266 of the other. The three water molecules, together with the terminal oxygens of the Y228 residues and all δ1 and ε2 nitrogens of the H266 residues, form a triangular hydrogen bonding network in the centre of the histidine bound ACT domain trimer, which is connected to the bound histidine. This network is absent in the ATP bound *Cje*ATP-PRT structure (Figure 2.25).

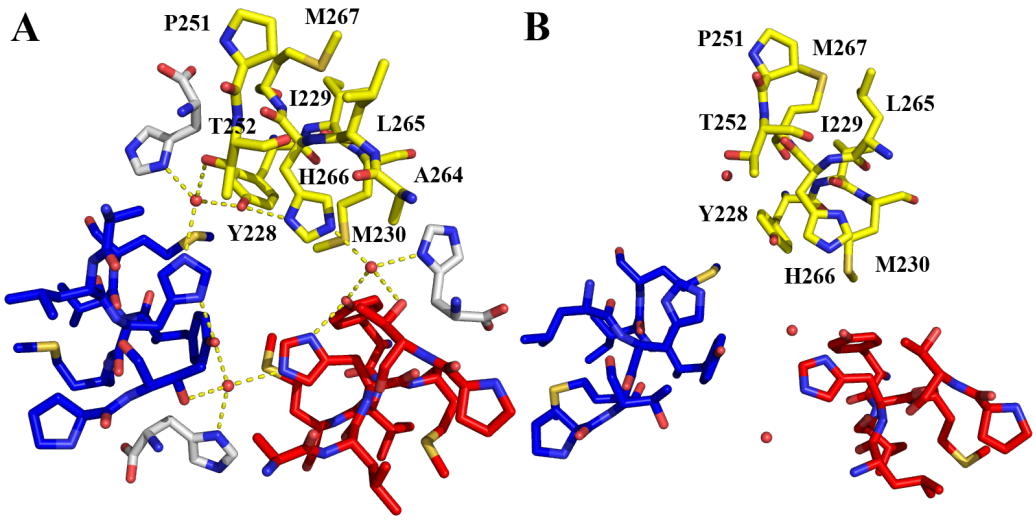


Figure 2.25: Hydrogen bonding network in histidine bound *Cje*ATP-PRT ACT trimer. Central residues of one ACT trimer of 4YB5 (A) and 4YB7 (B) in stick representation. The binding of histidine (white) causes the formation of a tight hydrogen bonding network (dashed lines) including three water molecules (small spheres), residues Y228, T252 and H266. Residue labels are placed on one chain only.

These interactions pull the two ACT surfaces together, which leads to the obvious compaction and rotation of the ACT domain trimers relative to the catalytic core in the inhibited conformation (Figure 2.24D). Similar observations have also been reported for *Mtu*ATP-PRT crystal structures (PDB codes: 1NH7 and 1NH8).[35]

The dimer interface also undergoes significant changes, which are most prominent between 4YB7 and 4YB5, where the calculated surface area is increased by approximately 200 Å. These differences are caused by a "flex" of the dimeric unit. In 4YB7 the interface is constituted predominantly by residue ranges T152–L163 (α6 and flanking loops) and T176–N180 (α7) of the first chain and Q37–I40 (α1-β2 loop), R8 (β1), and R54 and D57 (α2) of the second chain. In comparison to this, the two equivalent chains in 4YB5 are rotated relative to each other, retaining the contacts of α6 in the first chain to R8 and the α1-β2 loop in the second chain. The contacts between α7 and α1-β2 loop are lost entirely though, as α7 of both chains is shifted towards the outside of the hexamer. New close contacts are made instead

between residues N148–L151 (β8) of the first and D57–D64 (α2) of the second chain, which are now located in close proximity. The C-terminal ends of the α3 helices (L87) come significantly closer together (6.4 ± 0.3 Å vs. 13.8 ± 0.5 Å), which accounts in part for the large increase in surface area seen in 4YB5. The same rotational flex is observed in 4YB6 but to a lesser extent.

The neighbouring dimers of the inhibited hexamer also tighten their interfaces in an interlocking motion, clearly visible by the relative position change of the α7 helices in the inhibitor bound state compared to their diagonal side-by-side contact in the ATP bound structure (Figure 2.26).

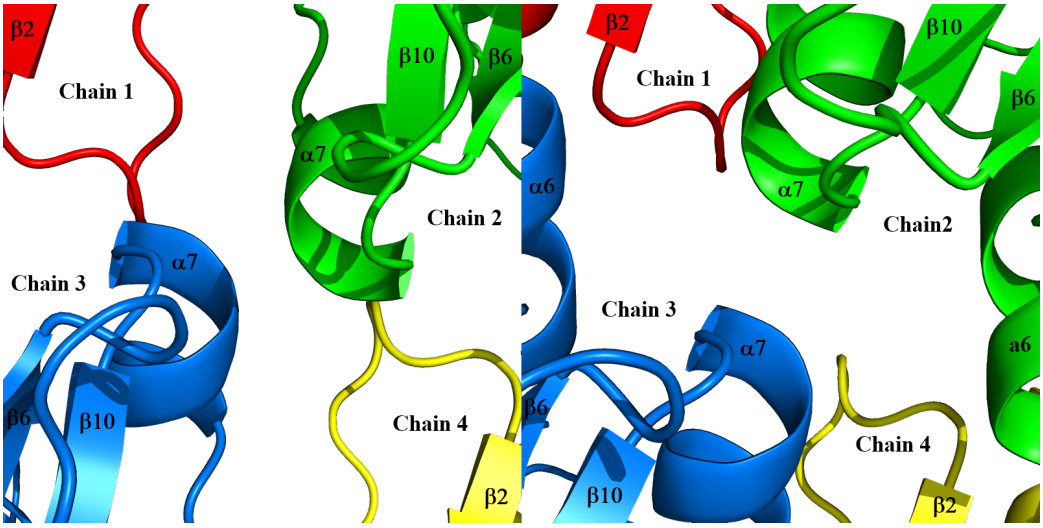


Figure 2.26: Comparison of α7 positions. The two different orientations of α7 at the contact point of two four neighbouring chains in the open (left hand side) and closed (right hand side) conformation of *Cje*ATP-PRT are displayed in the cartoon representation.

Situated on the inside of the hexamer, the active sites are directly influenced by the flex in the dimer interface. Following the twist of the central beta sheet of domain I, the β1-α1 and β3-α2 loops are turned away from the opposite chain. Or if examined from the perspective of domain I, the central beta sheet of domain II turns, retracting the α7 helix from the dimer interface. This is exemplified by the change in distance between the α carbon atoms of N180 and S38, 6.7 ± 0.5 Å in 4YB7 and 13.3 ± 0.3 Å in 4YB6, respectively.

This rearrangement of the interface affects the long basic side chain of the conserved residue K13. It is able to reach into the opposite active site in 4YB7, where K13 is suspected to interact with the pyrophosphate group of PRPP.[27,36] In 4YB6 K13 is significantly further away. Its α carbon distance to S154 is 10.7 $\pm$ 0.3 Å in 4YB6 compared to 7.7 $\pm$ 0.1 Å in 4YB7. On the other side of the active site the positioning of the α1-β2 loop is also altered. In 4YB6, this loop is found closing over the active site of the adjacent chain (not its dimer partner chain), coordinating the phosphate group of AMP sitting in the PRPP binding loop (β9-α7 loop), with the two histidine residues H33 and H35 in hydrogen bonding distance. In 4YB7 this "capping" of the active site loop does not occur, as the α1-β2 loop is far away from any possible interaction with the neighbouring subunit (Figure 2.27).

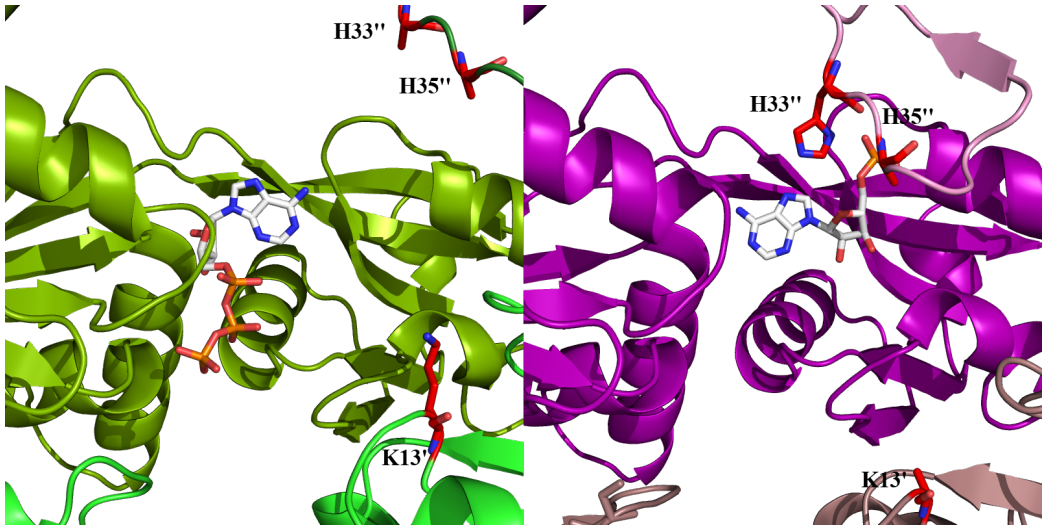


Figure 2.27: Changes in the *Cje*ATP-PRT active site between open and closed conformation. A hexamer of 4YB7 (LHS, green) and 4YB6 (RHS, purple) were superimposed on a single chain to compare the impact of the overall conformational changes in the enzyme complex. Three chains each are shown in the cartoon representation in variations of the ground colour with the superimposed chain in the centre. The bound active site ligand in shown in white. Residues K13', H33" and H35", representing the relative movement of structural motifs, are shown as red sticks. Heteroatoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

The active sites of 4YB5 and 4YB6 with all resolved side chains are close to identical (RMSD = 0.45 Å). The main difference between the two sites is the position of the α7 helix, which is shifted inwards (in respect to the hexamer surface) by approximately 4 Å in the AMP bound structure (4YB6). This is likely due to the presence of the 5'-phospho ribose moiety of AMP in the PRPP binding loop (β9-α7 loop) directly preceding the α7 helix. In the absence of AMP (4YB5), the PRPP binding loop possesses greater conformational freedom and allows the helix to move outwards, while it is held in place by the tight binding interactions of the PRPP loop in the presence of AMP (4YB6).

2.9.8 Comparison to existing crystal structures

The chain structure and overall fold observed for *Cje*ATP-PRT is identical to the folds described[26,35] in all deposited structures of ATP-PRT long form enzymes (PDB codes: 1NH7, 1NH8, 1H3D, 1Q1K, 2VD3). Due to the higher symmetry in those structures, symmetry operations must be applied to create hexamers of equivalent composition. As expected from the primary sequence, the structure of *Cje*ATP-PRT is most similar to the structure of *Eco*ATP-PRT, of which only the open form has been observed. Superimposition of *Eco*ATP-PRT chain (PDB code: 1H3D) with a single *Cje*ATP-PRT chain (PDB code: 4YB7) resulted in a close fit (RMSD = 0.988 Å), while the full length chain of *Mtu*ATP-PRT (PDB code: 1NH7) produced a weaker fit (RMSD = 2.056 Å). This is caused mainly by the different angle and orientation of the ACT domain in all three structures. If only the catalytic domains are considered, the superimposition of the two chains of *Cje*ATP-PRT (PDB code: 4YB7) and *Mtu*ATP-PRT (PDB code: 1NH7) consequently results in a better fit (RMSD = 1.216 Å) (Figure 2.28).

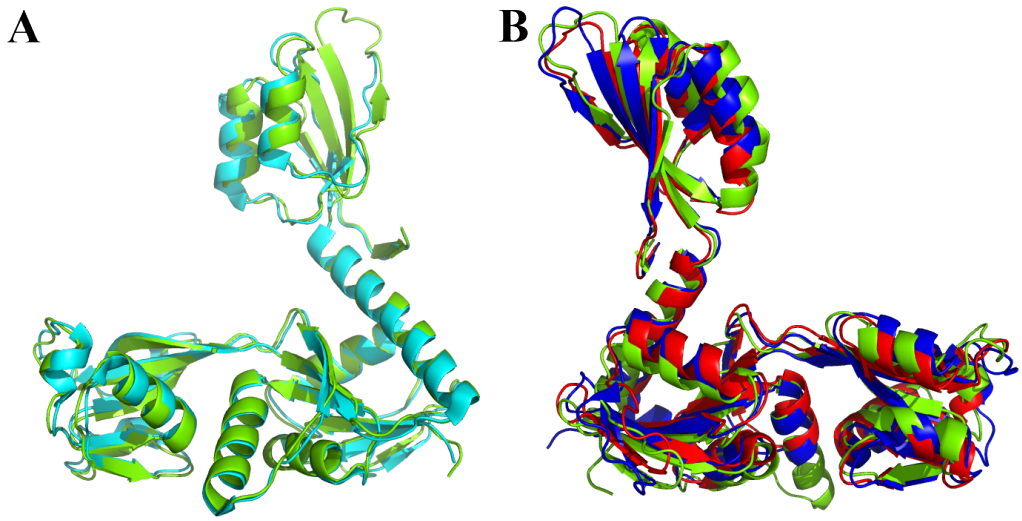


Figure 2.28: Comparison of ATP-PRT long form enzymes. Cartoon display of a single chain of all currently known long form ATP-PRT enzymes. A: Active conformation. Superposition of *Cje*ATP-PRT (4YB7) (green) with *Eco*ATP-PRT (1H3D)[26] (cyan). B: Inhibited conformation. Superposition of *Cje*ATP-PRT (4YB6) (green) with *Mtu*ATP-PRT (1NH8)[35] (red) and *Mth*ATP-PRT (2VD3) (blue).

Another difference is the much shorter helix $\alpha 3$ in *Mtu*ATP-PRT, which only includes eight residues compared to 14 residues for *Cje*ATP-PRT and *Eco*ATP-PRT. This difference can also be easily retraced to the sequence level (Figure 2.3). The same short helix was found in the inhibited form of the archaean *Mth*ATP-PRT (PDB code: 2VD3), which overlays similarly well (RMSD = 1.879 Å) with the inhibited *Cje*ATP-PRT (PDB code: 4YB6).

2.9.9 Small angle X-ray scattering

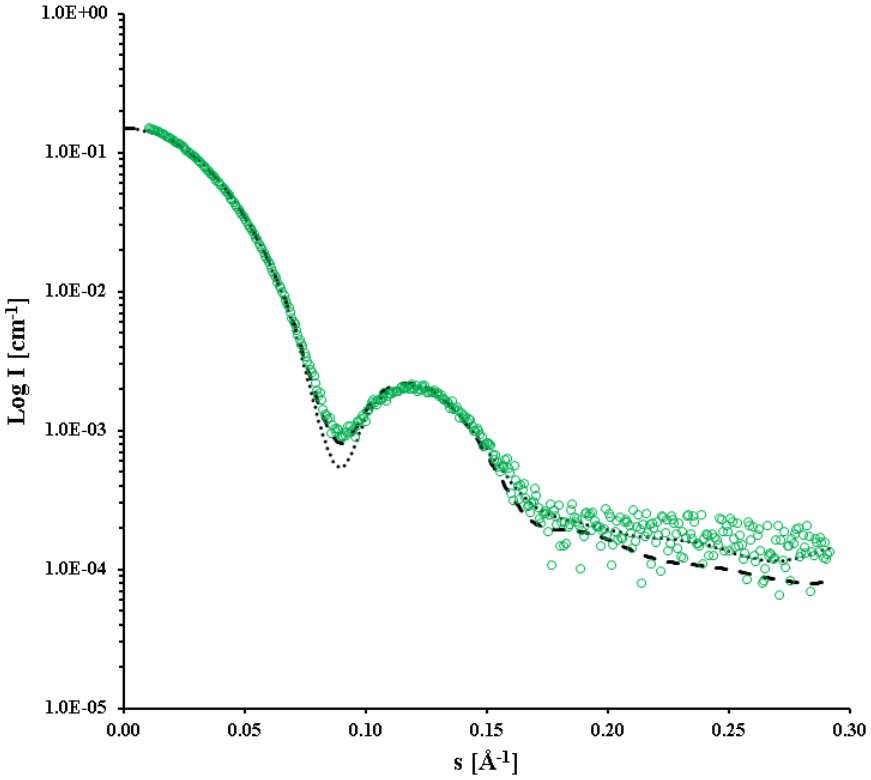
Small angle X-Ray scattering was employed to compare the active and inhibited crystal structures of *Cje*ATP-PRT with the shape of the enzyme in solution. Data was collected in the presence and absence of an inhibitory mix of 1 mM histidine and 2 mM AMP following elution from a SEC column (section 6.1.42).

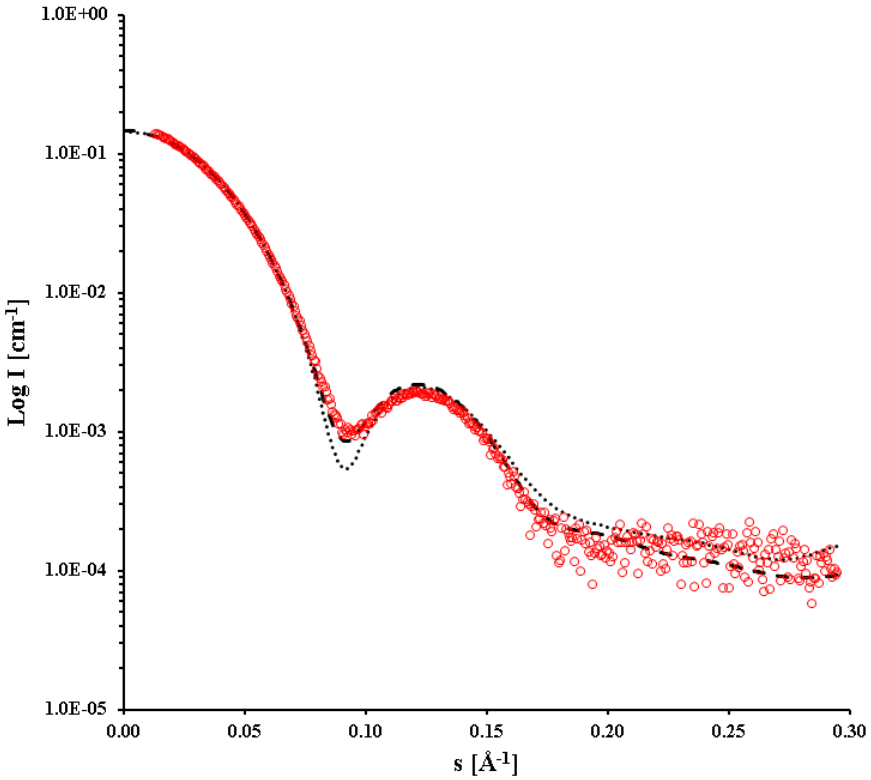
The radius of gyration (R_g), determined by the Guinier analysis, was 43.5 ± 0.4 Å for the scattering profile of the ligand-free dataset and 42.0 ± 0.4 Å in case of the inhibited enzyme complex. Using the indirect Fourier transform method[95] the pair distribution function (P(r)) was established with R_g being 43.5 ± 0.2 Å for the native and 42 ± 0.2 Å for the inhibited dataset. The maximum scattering particle (D_{max}) was calculated to 152 Å for the ligand-free and 147 Å for the inhibited enzyme data (Appendix D). Following this analysis the enzyme appeared almost unchanged in the uninhibited or inhibited form.

The obtained scattering profiles are consistent with a hexamer in solution as shown by the CRYSOL fits (Figure 2.29), with both data sets being nearly identical. The theoretical scattering of the ATP bound crystal structure (4YB7) shows marginal differences to the inhibitor bound structure (4YB6), which was not observed experimentally.

The overall fit quality was poor. The better fits were observed against the inhibitor-bound crystal structure with χ values of 4.4 (native) and 3.6 (inhibited). The fits against the ATP bound structure in turn produced χ values of 5.7 (native) and 4.2 (inhibited). This observation is in line with the previous results (section 2.8), suggesting that the only existing oligomeric species of *Cje*ATP-PRT is a homo-hexamer, but do not indicate a conformational change.

Figure 2.29 *(following page)*: SAXS measurements of *Cje*ATP-PRT. Theoretical scattering profiles of the *Cje*ATP-PRT crystal structures 4YB7 and 4YB6, representing the active (dotted line) and inhibited (dashed line) form of the hexamer, were generated by CRYSOL[96] and compared to the experimentally determined data in the absence (green) and presence (red) of 1 mM histidine and 2 mM AMP.


1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
Log I [cm-1]
0.00
0.05
0.10
0.15
0.20
0.25
0.30
s [Å-1]


1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
Log I [cm-1]
0.00
0.05
0.10
0.15
0.20
0.25
0.30
s [Å-1]

2.10 Investigation of *Mtu*ATP-PRT

The ATP-PRT enzyme from *M. tuberculosis* has been studied in great detail both kinetically[89] and crystallographically[35,75] and was therefore chosen to serve as a comparison to *Cje*ATP-PRT.

Soluble expression was achieved in pET24a(+). The *Mtu*ATP-PRT had been altered to possess a short N-terminal non-removable His-tag, which was used for purification via IMAC. The protein preparation was highly pure (Figure 2.30) but unfortunately found to be inactive in all tested reaction buffers tested (section 6.2.6).

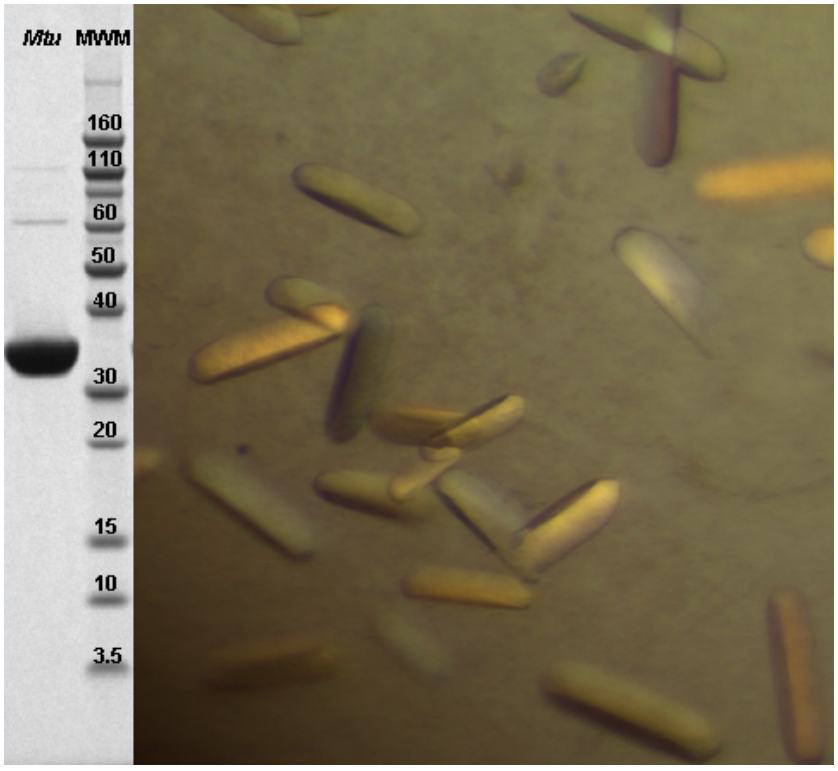


Figure 2.30: Crystallisation of *Mtu*ATP-PRT. LHS: Assessment of the purity of the *Mtu*ATP-PRT preparation. RHS: Photograph taken of drop containing *Mtu*ATP-PRT crystals. The crystallisation condition was 0.1 M MES pH 6.5, 1.8 M $MgSO_4$ (section 6.2.6).

Nonetheless, the protein was used for crystallisation making use of the previously described condition.[35] Small triangular prism-shaped crystals formed readily over night. ATP was soaked into the crystals and several diffraction data sets were collected at the Australian Synchrotron, one of which showed density for an active site ligand. The structure was solved by molecular replacement with the deposited ligand-free structure (PDB code: 1NH7) as starting model and revealed one ATP molecule bound in the catalytic site of the enzyme, which is a novel finding for *Mtu*ATP-PRT (Table 2.9). *Mtu*ATP-PRT crystallised in a trigonal spacegroup with only one chain in the unit cell, as in the previously published datasets 1NH7 and 1NH8. The overall shape of the chain was unsurprisingly more like 1NH7 (RMSD = 0.45 Å) than 1NH8 (RMSD = 1.27). Therefore it was compared in detail to the open, ATP bound, structure of *Cje*ATP-PRT (4YB7).

The two structures superimpose well with an overall RMSD value of 1.89 Å (Figure 2.31A). The main differences are, as described before, the shorter helix α3 and the missing β11-α7 and β13-β14 loops. The active site of *Mtu*ATP-PRT appears more compact (Figure 2.31B). The observed chain was leaning tightly against its dimeric counterpart, with the side chain of residue K9 extending into the active site of the other chain. There are three sulphate ions bound next to ATP in the active site that are interacting with the residues of the putative PRPP binding site, including the β9-α7 loop (residue range 157–161) and the conserved E141.

ATP adopts a similar binding mode in the active sites of *Mtu*ATP-PRT and *Cje*ATP-PRT, but the conformation and the interactions formed are different (Figure 2.32) because of the protein environment. *Cje*ATP-PRT possesses a coordination site for a single Mg^{2+} ion, which in turn binds to the triphosphate. In *Mtu*ATP-PRT, on the contrary, the functional Mg^{2+} is not coordinated by any protein side chains but bound to all three phosphates, giving the triphosphate group a different twist. A second Mg^{2+} ion is found shared between the two ATP molecules in a dimer. It was coordinated by the terminal oxygen atoms of both γ phosphates and both R49 side chains.

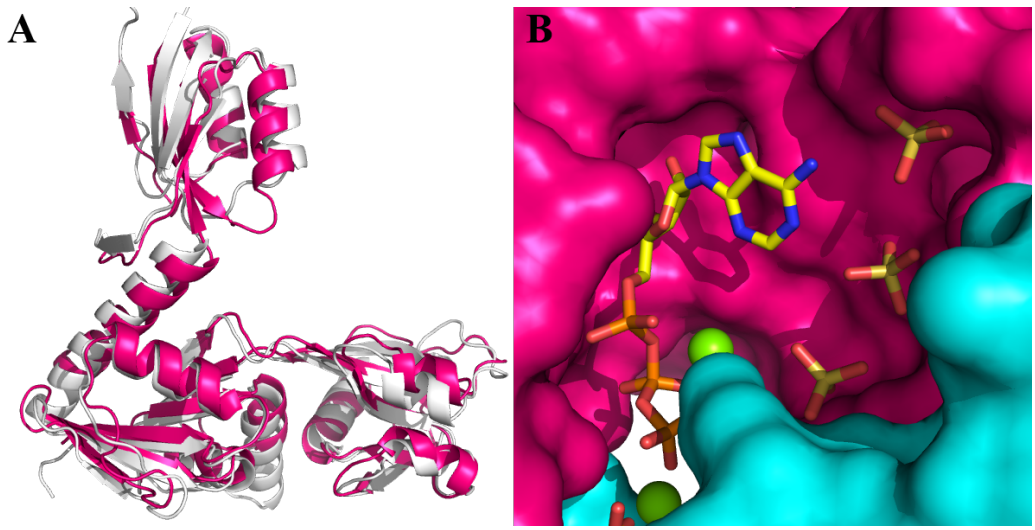


Figure 2.31: The *Mtu*ATP-PRT single chain and its active site.
A: Structural superimposition of the ATP bound *Mtu*ATP-PRT single chain (pink) with a single chain of *Cje*ATP-PRT (white) in cartoon representation. B: Surface representation of the active site of *Mtu*ATP-PRT with the bound ATP (bright yellow) and sulphur molecules displayed as sticks. Monomer contained in the unit cell at the top (pink) and its symmetry mate at the bottom (cyan). Atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), sulfur (yellow), magnesium (light green).

Additionally the interactions seen with *Cje*ATP-PRT R16 and Q12 do not exist in *Mtu*ATP-PRT. The only conserved interactions are with R54/R49, Asn75/Asp70 and L170/V155. Serine S191/S177 is another well conserved residue that forms hydrogen binding interactions to the ribose ring of ATP. In *Mtu*ATP-PRT this interaction is direct and not mediated by a water molecule. Overall, the two enzyme active sites bind ATP in a similar mode with a very limited number of essential interactions involving conserved or similar residues, while the rest of the interactions have evolved differently (Figure 2.32).

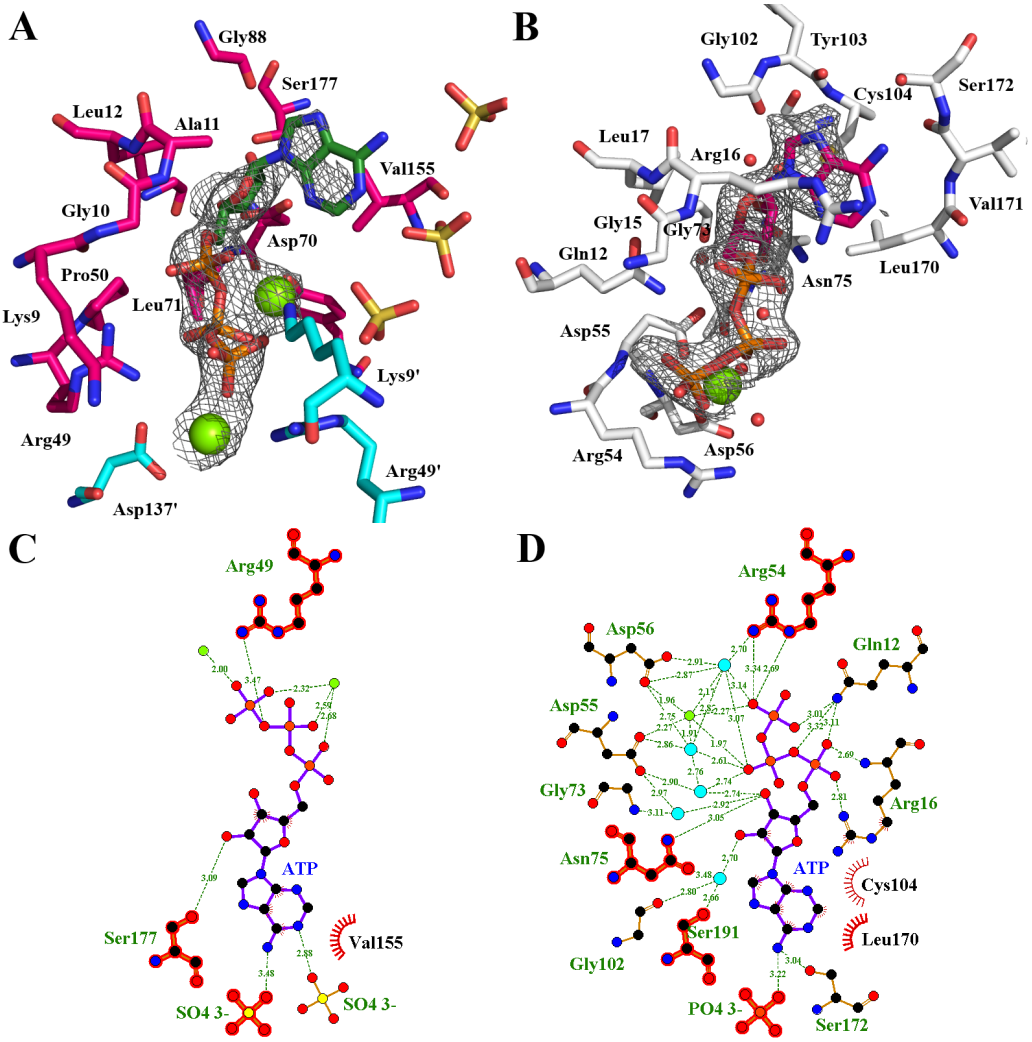


Figure 2.32: Comparison of ATP binding mode. Observed binding mode for ATP (dark green) in the active sites of *Mtu*ATP-PRT (pink) and *Cje*ATP-PRT (white) in a direct comparison. A and C: 3D stick representation of ATP with surrounding electron density (Fo-Fc map - grey mesh) and active site residues in close distance. The second chain of *Mtu*ATP-PRT is coloured in cyan. B and D: Planar plots of all ATP interactions in both structures were created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with the real space distance between atoms. Hydrophobic interactions are displayed as short radial red lines. Conserved interactions are highlighted in red. Water molecules are displayed as cyan spots. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), sulphur (yellow), magnesium (light green).

Table 2.9: Crystal parameters, data collection, and refinement statistics for *Mtu*ATP-PRT in complex with ATP.

	*Mtu*ATP-PRT
Data collection	
Crystal system	trigonal
Space group	$H3_2$
Unit cell parameters	
a, b, c [Å]	134.54, 134.54, 110.11
α, β, γ [°]	90.00, 90.00, 120.00
Resolution range [Å]	50.00–2.67 (2.81–2.67)
Measurements	68977
Unique reflections	10943
Completeness	99.8 (98.6)
Redundancy	6.3 (6.3)
$I/\sigma(I)$	15.7 (2.6)
R_{merge}	0.106 (0.805)
$CC_{1/2}$	0.997 (0.793)
Wilson B value [Å^2]	50.5
Matthews coefficient	3.05
Refinement	
Resolution [Å]	40.00–2.67
R_{cryst}	0.220
R_{free}	0.281
Chain length	293
Observed residues	272
Water molecules	0
Others	1 ATP, 2 Mg^{2+}, 8 SO_4^{3-}
Mean B [Å^2]	
Protein	49.13
Other	67.32
ATP	61.87
RMSD	
Bond lengths [Å]	0.007
Bond angles [°]	1.303
Chiral volumes [Å^3]	0.067
Ramachandran	
Preferred %	98.69
Allowed %	1.31
Outliers %	0.00
PDB entry	4YG9

2.11 Discussion

*Cje*ATP-PRT has been successfully cloned and purified using heterologous expression in *E. coli*. An unusual feature was the intrinsic affinity of the native *Cje*ATP-PRT enzyme to immobilised metal ions that made reverse IMAC purification impossible. This effect may involve the α1-β2 loop, which houses two histidine residues (H33 and H35) in a good configuration to chelate ions. In the active, ATP bound, structure this loop is found at the periphery of the main opening into the central cavity, while it is close to the active site (β9-α7 loop) in the inhibited conformation. This suggests that conformational flexibility in the active form opens up the possibility for the histidine side chains to bind particles at the outside of the hexamer. This effect could potentially be circumvented by purifying the enzyme in the presence of low concentrations of histidine ($\leq$ 1 mM), which fully inhibit the enzyme but do not reduce the IMAC binding affinity to the His-tag.

Overall, the kinetic parameters determined for *Cje*ATP-PRT were in agreement with those of other long form ATP-PRTs. Although the ATP-PRT enzymatic reaction is well studied, extensive kinetic studies are rare. It was therefore unsurprising that the acquired K_M and k_{cat} values were closest to the data published for *S.* Typhimurium, which was grouped in the same clade of the protein sequence alignments as *Cje*ATP-PRT, even though the work was performed using partially pure enzyme preparations.[28] The recently published kinetic data for *Mtu*ATP-PRT reports 3-fold higher K_M values than measured for *Cje*ATP-PRT and can still be seen as a good comparison.[89] Unfortunately the *Mtu*ATP-PRT preparation presented in this thesis was inactive and could therefore not be compared. The inhibitory behaviour of AMP was clearly competitive towards both substrates which is in accordance with previous findings. The determined K_i values are also very close to *Sen*ATP-PRT data (0.48 and 0.55 mM against ATP and PRPP respectively).[18] The competitive behaviour is evident by the binding modes of AMP and ATP presented in the obtained crystal structures (4YB6 and 4YB7). AMP bridges over both binding sites covering the adenine ring po-

sition of ATP, as well as the phosphate coordination site of PRPP in the highly conserved PRPP binding motif, as has been previously described for *Eco*ATP-PRT.[26] The inhibition mode of histidine was found to fit best to a non-competitive model against both substrates, contrary to previous descriptions of the histidine vs. ATP inhibition mode as uncompetitive.

The kinetic results were supported and largely confirmed by the affinity constants measured with ITC, adding a new level of understanding and confidence to the analysis. The binding curves followed a single site model for all used ligands with the exception of histidine. Since this is the first description of ITC experiments on an ATP-PRT enzyme, no comparison is available and data interpretation is challenging. The complex curvature of the histidine binding experiments clearly indicated a cooperative effect leading to the assumption that the binding of the first histidine molecule to a hexamer influences all other binding sites in the complex. It was therefore fitted with a sequential binding model in all cases. All resulting values should be interpreted with care as the initial strong exothermic signal in the solely histidine binding experiment was not observed when the enzyme was pre-incubated with AMP. It appears possible that this strong exothermic event is based on the rearrangement of the enzyme complex upon histidine binding, adopting the inhibited state with close contacts at the ACT domain interfaces. It has been shown for *Eco*ATP-PRT that AMP binding alone does not cause the complex to change into the inhibited form,[26] which could explain why this effect is not seen for AMP titrations. On the other hand it is also absent when the enzyme has been pre-incubated with AMP, suggesting that AMP has an influence on the conformation of the complex or at least on the energy required for the transition.

The determined K_d value for AMP was significantly lower than the kinetically measured competitive K_i values, as well as the reported K_i values for *Sen*ATP-PRT. By measuring the binding of AMP to enzyme in the presence of kinetically relevant concentrations of ATP (1 mM), it was shown that this discrepancy is dependant on the ATP concentration in the ITC experiment. Without the substrate competing for the active site, a tighter

binding was observed than in the presence of 1 mM ATP. This conclusion could apply equally to the differences of the measured substrate K_d and K_M values, which occurred most likely due to the absence of second substrate.

It was shown evidently that the kinetic mechanism of ATP and PRPP to *Cje*ATP-PRT follows a sequential model. Considering the difference in the determined K_M values and the relative instability of PRPP compared to ATP it appears likely that ATP binds first, preparing the active site for PRPP binding, which is then quickly turned over. This hypothesis is also supported by the product inhibition and reverse reaction studies performed on *Sen*ATP-PRT.[17,28] However ITC experiments clearly showed that binding of PRPP is possible in the absence of ATP, although the affinity is low. Together with the observation that the ATP K_d was lower than the determined K_M for ATP (about half), this suggests that both substrate molecules are able to bind the empty active site. It is therefore proposed that the binding mode of *Cje*ATP-PRT is random sequential. This is further supported by the observed stabilisation effect of *Cje*ATP-PRT on PRPP under assay conditions. It is plausible that PRPP adopts different binding modes in the absence and presence of ATP, enabling the enzyme to secure this energetically expensive and labile metabolite without the risk of spontaneous decomposition, until the more stable substrate ATP binds and the reaction occurs.

It has been discussed that ATP-PRT long form enzymes exist in an equilibrium between an active dimer and an inactive hexamer, which is driven by protein concentration and the presence of inhibitor molecules. All data presented in this chapter is consistent with a single hexameric species under all examined conditions, including kinetically relevant enzyme concentrations, posing the question whether or not an equilibrium exists at all or if the homo-hexamer is the only active and biologically relevant species. One of the main arguments for the existence of the dimer-hexamer equilibrium is the fact that all active sites lie on the inside of the hexameric arrangement, apparently inaccessible to any substrate molecule. A comparison of the dimer and trimer interfaces of both *Mtu*ATP-PRT and *Eco*ATP-PRT revealed that

the dimer interface is stronger and more important for complexation, rendering the trimer interface auxiliary or even optional. This was further supported by the comparison to the short form enzymes, which contribute two ATP-PRT dimers to the overall hetero-octameric complex highlighting the importance of the dimer.[37] The PISA[93] analysis of the three *Cje*ATP-PRT crystal structures delivered a different result. The trimer interfaces of the *Cje*ATP-PRT ACT domains were found to be of equal or more importance than the dimer interfaces, particularly in the absence of histidine, indicating that the hexameric species is likely to be stable.

Furthermore the hexameric structures of all structurally characterised ATP-PRT long form enzymes possess the same large openings between neighbouring domains, which would likely permit access to the interior of the complex at any time. Another possible entrance for substrate molecules is present in the form of a small opening in the dimer interface, observed in both ATP bound structures (4YB7 and 4YG9), that leads directly to the active site and may function as a substrate channel. In the same way, the hexamer could loosen and widen in a "breathing" motion, allowing substrate access through the weak interfaces between the catalytical dimer cores. Limited evidence for such movement comes from the thermal stability experiments performed on *Cje*ATP-PRT. Firstly the synergistically inhibited enzyme is significantly more stable than the unliganded enzyme (thermal shift of 5°), which is commonly associated with increased rigidity. This observation goes hand in hand with the crystallographic finding that the closed hexamer form is narrower, tighter and generally more refined than the open form. This increased rigidity was perhaps further evidenced by the findings of the DSF experiments, when the fluorescent dye exhibited premature binding to the ligand-free enzyme, but did not in the presence of both AMP and histidine. The prior discussed metal affinity and the low symmetry of the APT bound crystal structure also support the conclusion that the uninhibited hexamer possesses a certain flexibility, allowing for substrate molecules to enter.

The observed changes between the open and closed conformations of the *Cje*ATP-PRT hexamer are a fine interplay between changes in the in-

terfaces and the resulting shifts of functional structural motifs. The clearest example here is the residue range G29–P47, which forms a long loop including β2. While its C-terminal side is involved in the dimer interface formation in the active hexamer, it is largely released upon histidine binding and allowed to move inwards, where it then caps the phosphate binding site of a neighbouring subunit with its N-terminal side. These rearrangements of structure motifs between neighbouring chains are an additional indication for the importance of the hexameric state for the function of *Cje*ATP-PRT.

Although the most significant changes in the hexameric complex are seen upon histidine binding, the two crystal structures 4YB5 (histidine only) and 4YB6 (AMP and histidine) also show small differences. The described active site flex is less prominent in 4YB6, resulting in a smaller calculated dimer interface area, while the PRPP binding loop and helix α7 are significantly more flexible in 4YB5. These differences may be attributed to the packing forces present in the alternative crystallisation condition. Alternatively, these changes may be the result of the binding of AMP in the active site. In the absence of active site ligands, histidine binding forces the hexamer to adopt a very tight conformation, allowing only small functional motifs to move freely. The binding of AMP, on the other hand, could release some of the strain by reorienting the PRPP binding loop, causing the dimer interface to relax. These small differences could potentially be part of the synergistic effect of histidine and AMP inhibition, although the conditions presented in the structures should both represent an absolutely inhibited enzyme, due to the high inhibitor concentrations used.

The structural information presented in this chapter provides great detail about the binding modes of two of the ATP-PRT ligands in the active site. Convincing electron density for both ATP and AMP was found, which led to the analysis of their binding interactions. PRPP binding in the active site has not been shown experimentally. It is expected to occur in a similar manner to the well established binding in type I and II PRTases,[22] heavily involving the residues of the β9-α7 loop assisted by an absolutely conserved glutamate residue (*Cje*ATP-PRT E156). The reported ATP binding, as seen

in *Cje*ATP-PRT as well as in *Mtu*ATP-PRT, has not been described previously. This enabled a detailed comparison of the binding sites in light of this new ligand, highlighting the different architectures of both active sites with only a few strictly conserved interactions: R49/R54 binding the triphosphate, D70/N75 and S177/S191 forming contacts to the ribose oxygens and V155/L170 interacting with the adenine.

As the *Mtu*ATP-PRT protein was found to be inactive, it remains unclear whether the solved structure shows the correct orientation of ATP or is merely an artefact of crystallisation. The high precipitant concentrations could have allowed for the active sites to open up for ligand binding while they were permanently blocked or misfolded in solution as a consequence of purification. As no further experiments have been performed on this enzyme, it could also be possible that no suitable kinetic condition was found, although the preparation was potentially active. The lack of activity may alternatively be the result of an inability to bind PRPP. Overall it seems unlikely for a large molecule like ATP to adopt a false conformation and interactions, given the similarity to the ATP binding mode observed in *Cje*ATP-PRT.

Significant differences were observed between the AMP binding mode presented in crystal structure 4YB6 and its counterpart in *Mtu*ATP-PRT (PDB code: 1NH8). Due to the absence of a large side chain (R16 in *Cje*ATP-PRT) "capping" the adenine binding site, the bound AMP possesses more rotational freedom in *Mtu*ATP-PRT. AMP is able to flip upside down to interact freely with the side chains of D30 and D33 of the α1-β2 loop, while the longer side chains H33 and H35 in *Cje*ATP-PRT coordinate the 5'-phosphate bound in the β9-α7 loop. Although involving different residues and connectivity both interactions have the same overall effect and could be seen as an example of fine tuning of an enzyme active site by parallel evolution. Alternatively it is possible that the density for AMP in *Mtu*ATP-PRT[35] was misinterpreted and not sufficient for the placement presented in 1NH8. This would leave the binding mode of the AMP ribose as observed in *Eco*ATP-PRT (PDB code: 1H3D)[26] as the only point of comparison, which is nearly identical to the one presented for *Cje*ATP-PRT.

There was no observed difference between the SAXS profiles in the presence and absence of the inhibitors mix. The maximum particle size and radius of gyration calculated from the uninhibited data set were slightly larger than in the data set containing the inhibitors. This is in accordance with the tightening of the crystal structure in the presence of histidine. But both data sets fitted better to the inhibitor bound X-ray crystal structure although the expected scattering for the substrate bound structure was marginally different. As the SAXS experiments have not been performed in the presence of ATP it can also not be ruled out that the substrate has a determining effect on the hexameric structure in solution, promoting the occurrence of the expected open form. Aggregation might also have negatively affected the outcome of the experiments and data recollection with an unfrozen enzyme sample should be considered.

Chapter 3

Biochemical and structural characterisation of the *Campylobacter jejuni* adenosine triphosphate phosphoribosyltransferase Core mutant

3.1 Introduction

ATP-PRT enzymes are divided into members of the long or the short form. The previous chapter described the detailed analysis of a new member of the long form ATP-PRT from *C. jejuni*. This enzyme possesses a C-terminal regulatory domain characteristic of the long form, which is classified as an ACT domain by fold and function.[41] The *Cje*ATP-PRT ACT domain facilitates histidine binding and inhibition of the phosphoribosyltransferase reaction resulting in a conformational change of the homo-hexameric enzyme com-

plex. The short form ATP-PRT enzymes do not possess this mechanistic feature. Instead a separate protein, HisZ, is employed for regulatory and catalytic purposes, leading to the formation of a hetero-octameric complex. The mechanism by which these complexes function has not been elucidated in detail to date, but the central feature is thought to be the alteration of the properties of the catalytic dimers[36,97] by similar movements to those observed in *Cje*ATP-PRT and other long form enzymes.

Since the architecture of the catalytic domains of long and short form ATP-PRTs is very similar and most important active site residues are conserved, it is likely that they have evolved from the same precursor. Whether the short form complex of contemporary proteins represents this precursor is unclear, but it is likely that the long forms evolved subsequently from an event that recruited an ACT domain to an ancestral ATP-PRT.[37] This chapter describes the creation of a truncated *Cje*ATP-PRT mutant reversing this evolutionary event on a modern long form ATP-PRT to study its properties as an isolated catalytic domain species.

3.2 Selection of the truncation site

The choice of the truncation position is crucial to the outcome of a large scale manipulation of a protein. For the removal of the regulatory ACT domain of *Cje*ATP-PRT the short α8-β12 loop (Figure 3.1) was chosen as it represents the hinge by which the ACT-domain movement is facilitated. The presence of exclusively polar residues made this position an excellent choice, considering the properties of the newly created C-terminus. The minimal exposure of new surface area to the environment, on the mainly polar outward face of the α8 helix, was also an important aspect in its favour. The exact chain termination position was chosen after residue E225, completely removing the residue range 226–299.

The mutant was named *Campylobacter jejuni* adenosine triphosphate phosphoribosyltransferase Core mutant (*Cje*ATP-PRT Core or simply Core

mutant), as it only possesses the catalytic domains I and II, which are positioned in the "core" of the *Cje*ATP-PRT wild type hexamer.

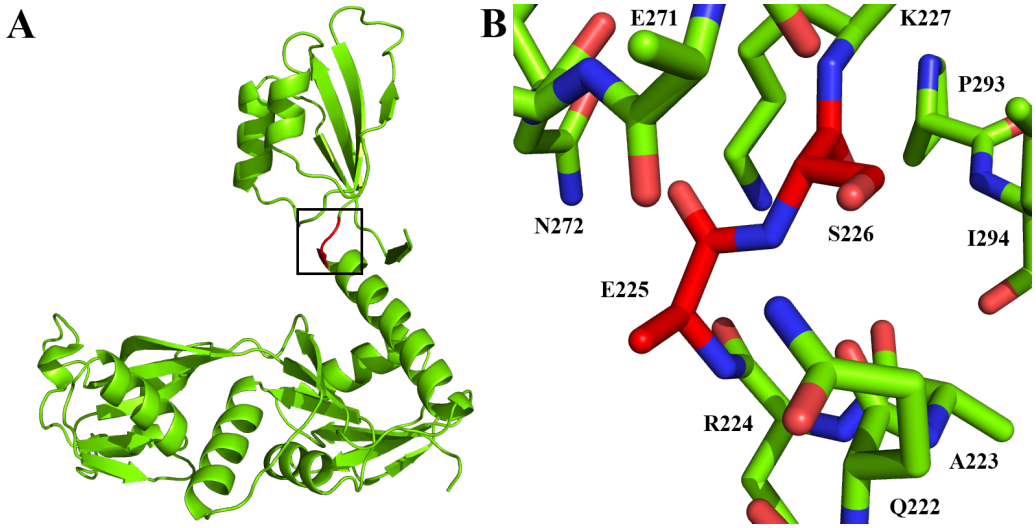


Figure 3.1: Truncation site. The truncation site (bright red) on the full length *Cje*ATP-PRT chain (green) is shown. A: Cartoon representation of the entire chain. The square box indicates the zoomed in area. B: Zoom on the truncation site. All residues are displayed in the sticks representation. Atoms are coloured according to element: oxygen (red), nitrogen (blue).

3.3 Cloning and purification

Cloning of the truncated mutant was performed following the procedure described for the *Cje*ATP-PRT wild type, but made use of an alternative reverse primer (section 6.3.1) that introduced the artificial translation stop (Figure 3.2). Initial screens of the pDEST™17 clone in BL21* (DE3) or Chaperone3 cells showed very limited solubility of the overexpressed protein. The gene was therefore transferred into pDEST™15, which adds a glutathione sulfur-transferase (GST) solubility tag. This successfully increased the solubility of the protein chain in BL21* (DE3) cells, which were grown at 37 °C for 4 hours after induction with 0.5 mM IPTG. The crude lysate was found to be active under standard *Cje*ATP-PRT assay conditions. Therefore these conditions were used for all large scale expressions of *Cje*ATP-PRT Core.

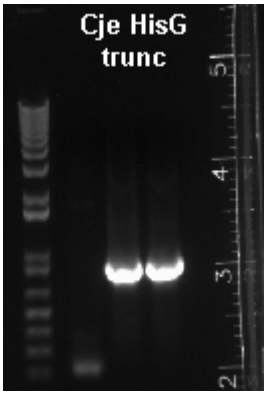


Figure 3.2: Cloning of *Cje*ATP-PRT Core. Agarose gel showing the amplified gene encoding for *Cje*ATP-PRT Core from three selected clones, using the colony PCR set up described in sections and 6.1.17. The gel was run according to section 6.1.13.

The *Cje*ATP-PRT Core mutant was observed to possess the same intrinsic affinity to IMAC columns. This fact was exploited in the development of the purification protocol. Cells were harvested and lysed according to the protocol described for the ATP-PRT wild type enzyme. An initial IMAC run isolated the overexpressed enzyme from the soluble crude lysate. The enzyme was then treated with TEV protease for 1 h at 37 °C and 180 rpm, which removed the tag completely without adversely affecting the solubility of the *Cje*ATP-PRT Core. TEV protease treatment was followed by a reverse purification step. Using GST affinity chromatography the free GST was removed from the protein preparation, while the *Cje*ATP-PRT Core eluted in the flowthrough. SEC was employed as the final purification step, removing the TEV protease and desalting the enzyme. After final assessment of the protein purity by SDS-PAGE (Figure 3.3), the *Cje*ATP-PRT Core containing fractions were concentrated to a stock concentration of $\geq$ 20 mg mL $^{-1}$ and kept at 4 °C for immediate use, or flash frozen in liquid nitrogen and stored at -80 °C.

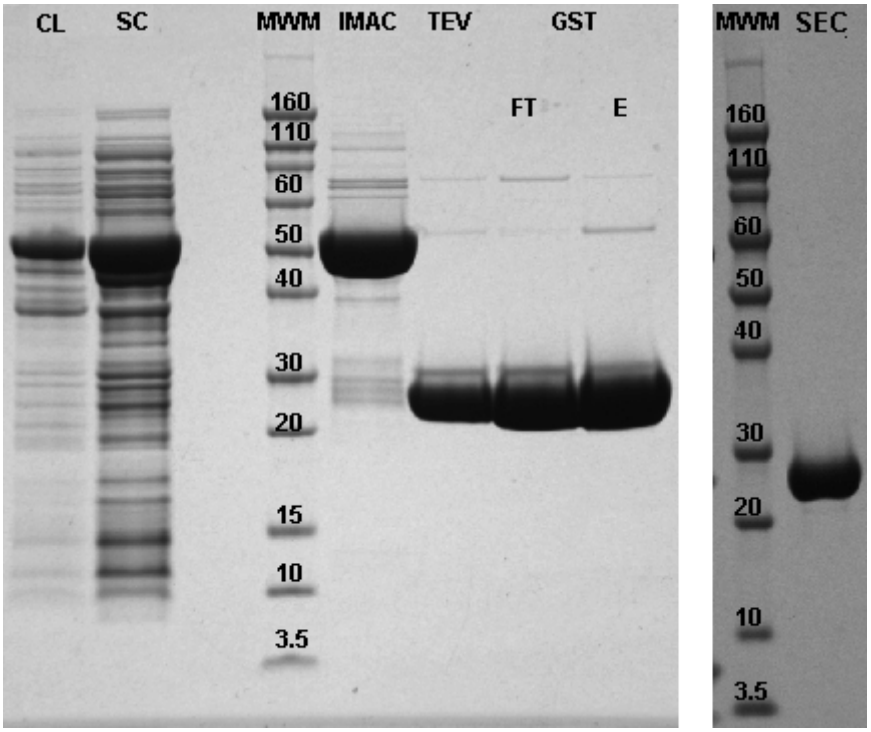


Figure 3.3: Purification of *Cje*ATP-PRT Core. SDS-PAGE showing the enzyme preparation at its various purification steps. Electrophoresis was performed according to section 6.1.31 in 10 % Bis-Tris gels using MES buffer. The theoretical molecular weight of *Cje*ATP-PRT Core is 53.9 kDa (tagged) and 25.3 kDa (tag removed). Lane labels correspond to: CL - crude lysate, SC - soluble crude, MWM - molecular weight marker, IMAC - eluate from IMAC, TEV - TEV protease treated, GST FT/E - flow through/eluate from GST affinity chromatography, SEC - Eluate from SEC.

3.4 Determination of molecular weight

The molecular mass of the purified *Cje*ATP-PRT Core was determined to be 25307.7 g mol^{-1} by electrospray ionisation mass spectrometry. This value is in close agreement with the theoretical value of 25308.2 g mol^{-1} calculated for the 226 amino acid sequence derived from the first 225 amino acid residues of the *Cje*ATP-PRT open reading frame and a single N-terminal glycine residue left over from the truncated tag.

3.5 Oligomeric state analysis

3.5.1 Size exclusion chromatography

During the purification process *Cje*ATP-PRT Core exhibited a different behaviour to the wild type enzyme. The relative ease with which the GST tag was removed and a comparatively late elution of the protein from the size exclusion column were indicative of a change in assembly or shape of the truncated mutant compared to the *Cje*ATP-PRT wild type.

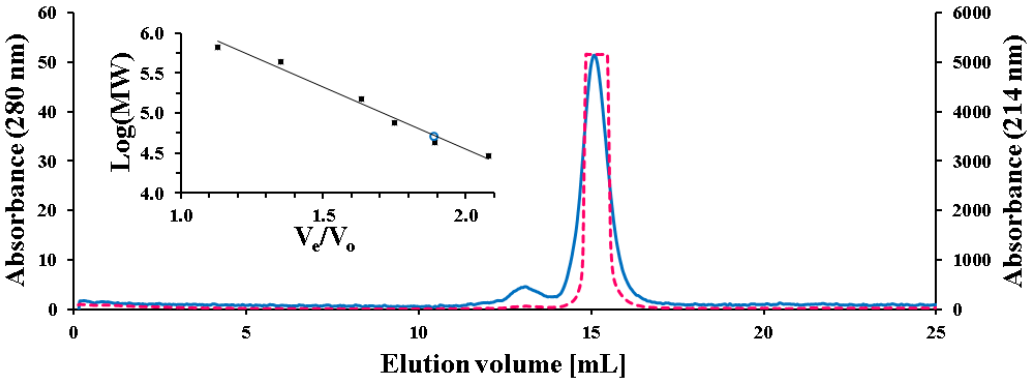


Figure 3.4: Analytical size exclusion chromatography of *Cje*ATP-PRT Core. SEC was performed on 1 mg ml $^{-1}$ *Cje*ATP-PRT Core using a Superdex^{TM} 200 10/300 GL column as described in section 6.1.26. The absorbance of the protein sample was monitored at 280 nm (blue, solid) and 214 nm (pink, dashed) due to the low extinction coefficient of *Cje*ATP-PRT Core at 280 nm. Inset: Calibration curve. The molecular weight of the standards proteins (Log(MW)) is plotted on a logarithmic scale against their normalised elution volumes (V_E/V_0). The void volume (V_0) was determined using Blue Dextran. The calculated molecular weight of the *Cje*ATP-PRT Core species is displayed as an open circle.

Analogous to the analysis of the wild type enzyme, analytical size exclusion was employed to estimate the molecular weight of the purified *Cje*ATP-PRT species at a concentration of 1 mg mL^{-1}. The main peak eluted from the column after 15.1 mL of isocratic flow (Figure 3.4). This peak had an expected mass of 52 kDa, which was very close to the theoretical dimer mass of 51 kDa. A second much smaller peak eluted at 13.1 mL corresponding to a size of 126 kDa, which may correspond to a pentamer or possibly hexamer.

3.5.2 Static light scattering

Analysis of the *Cje*ATP-PRT Core by SLS (section 6.1.40) was consistent with the presence of a dimeric species. Only a single peak corresponding to a single protein species with the molecular mass of 49 kDa was observed.

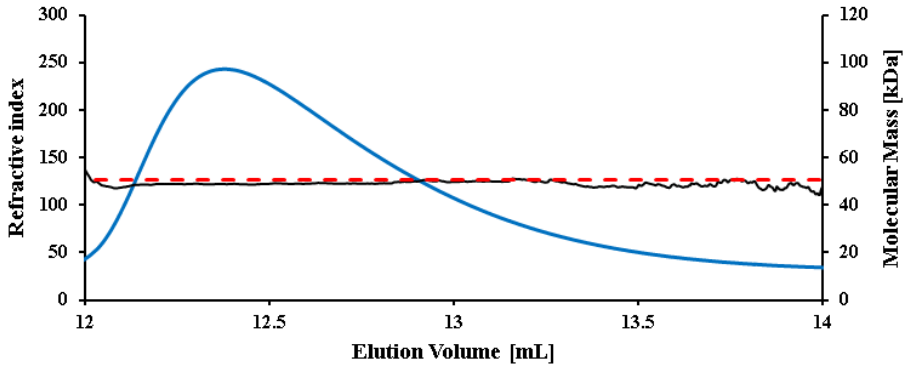


Figure 3.5: Static light scattering of *Cje*ATP-PRT Core. The refractive index of 1 mg mL $^{-1}$ *Cje*ATP-PRT Core (solid blue line) was determined as described in section 6.1.40. The molecular weight of the mono-disperse species (solid black line) eluting between 12 and 14 mL was calculated, using bovine serum albumine (BSA) as standard, and compared to the theoretical molecular mass of the *Cje*ATP-PRT Core dimer (red dashed line).

3.5.3 Analytical ultracentrifugation

Sedimentation velocity experiments further supported the findings of SEC and SLS. The main species detected had a sedimentation coefficient of 3.4, corresponding to a molecular mass of 49.8 kDa (Figure 3.6). A secondary peak at 5.4 S corresponding to a mass of 158 kDa was also observed with an increased prominence at the highest tested concentration (1.2 mg mL $^{-1}$). This may correspond to a small amount of a hexameric species (theoretical mass of 152 kDa) or a non-functional aggregate.

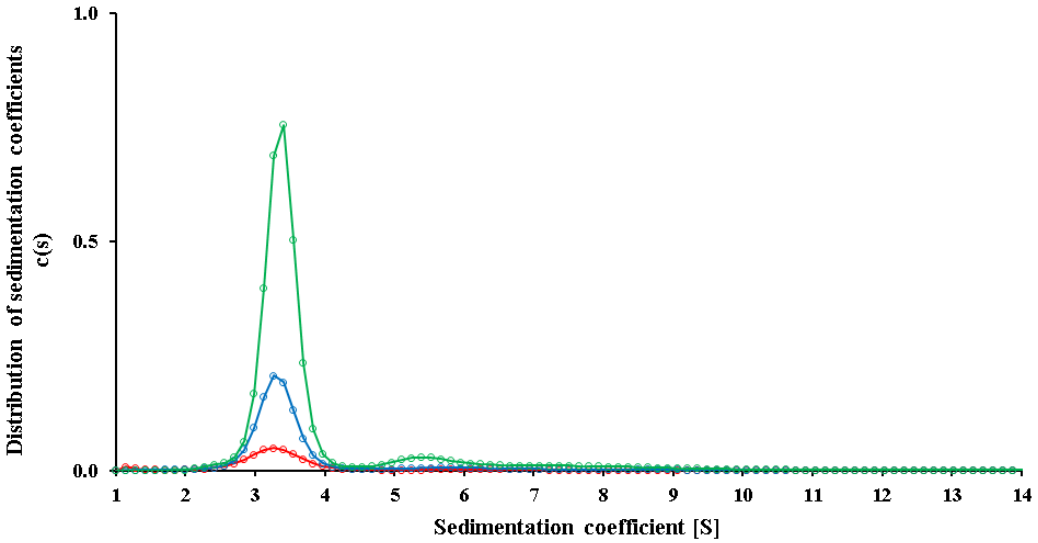


Figure 3.6: Size distribution function of *Cje*ATP-PRT Core. AUC experiments were conducted according to section 6.1.41. The resulting data was analysed using SEDFIT.[90] The plot shows the data of all three used enzyme concentrations un-normalised from highest to lowest signal: 1.2 mg mL $^{-1}$ (green), 0.4 mg mL $^{-1}$ (blue) and 0.13 mg mL $^{-1}$ (red).

3.5.4 Small angle X-ray scattering

Small angle X-ray scattering was also used to investigate the in solution behaviour of *Cje*ATP-PRT Core. Static solution data was collected in the presence and absence of 2 mM AMP using the capillary set up (section 6.1.42). The radius of gyration (R_g), determined by Guinier analysis, was 25.8 ± 0.1 Å for the scattering profile of the ligand free dataset and 24.8 ± 0.1 Å in the case of the AMP containing sample. Using the indirect Fourier transform method[95] the pair distribution function (P(r)) was established with R_g being 26.2 ± 0.1 Å for the unliganded and 24.5 ± 0.1 Å for the AMP containing dataset. The maximum scattering particle (D_{max}) was calculated for native and AMP containing data to 91.3 Å and 80.8 Å respectively. Overall data sets appeared fairly similar showing no significant change in assembly.

The scattering profiles were compared to the theoretical scattering of the PR-ATP bound *Cje*ATP-PRT Core crystal structure (section 3.8), which consists of two chains arranged as a dimer. Both datasets produced close

fits against this crystal structure with χ values of 1.3 (ligand-free) and 0.4 (containing AMP). A comparison was also made to the data acquired for the *Cje*ATP-PRT wild type, showing the vast differences in scattering behaviour between the two proteins. Fitting of the wild type data to the Core crystal structure using CRYSOL[96] resulted in χ values > 30, as would be expected. These results are consistent with the presence of the *Cje*ATP-PRT as a dimer and support all previous results.

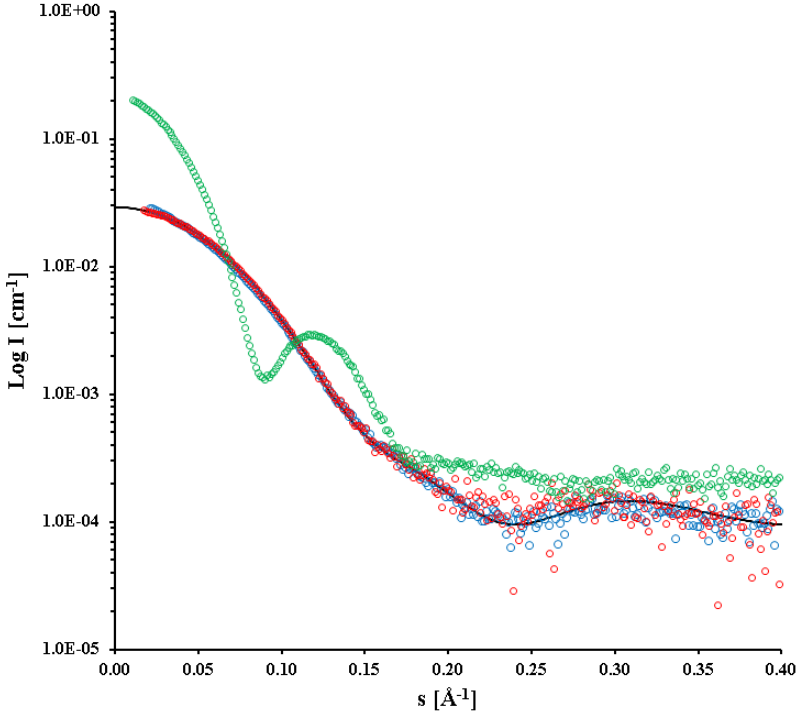


Figure 3.7: SAXS measurements of *Cje*ATP-PRT Core. The theoretical scattering profile of the PR-ATP bound *Cje*ATP-PRT Core crystal structure (solid black line) was generated by CRYSOL[96] and compared to the experimentally determined data of *Cje*ATP-PRT Core in the absence (blue) and presence (red) of 2 mM AMP, and *Cje*ATP-PRT wild type (green).

3.6 Kinetic properties

The kinetic parameters of the *Cje*ATP-PRT Core mutant were determined under the conditions stated for the wild type enzyme (sections 6.1.33 and 6.3.3). The K_M(ATP) value of *Cje*ATP-PRT Core is comparable to the *Cje*ATP-PRT wild type, whereas the K_M(PRPP) value increased about 3-fold (Table 3.1) compared to the wild type enzyme. This suggests that the substrate affinities of wild type and Core mutant are largely comparable. The turn over rate on the other hand was heavily affected with a difference of two orders of magnitude. *Cje*ATP-PRT Core is therefore a very inefficient catalyst.

Table 3.1: Kinetic constants of *Cje*ATP-PRT Core. Kinetic measurements were performed according to section 6.3.3. Equivalent values for *Cje*ATP-PRT wild type are given as a comparison.

Parameter	*Cje*ATP-PRT Core	*Cje*ATP-PRT wild type
K_M(ATP) [μM]	94 ± 11	97 ± 10
K_M(PRPP) [μM]	41 ± 5	15 ± 1
k_{cat} [s^{-1}]	0.024 ± 0.003	1.8 ± 0.1
IC_{50}(histidine) [μM]	no inhibition	44 ± 1
IC_{50}(AMP) [mM]	0.42 ± 0.05	2.1 ± 0.3

Due to the lack of the regulatory domain, the binding site for histidine is absent in *Cje*ATP-PRT Core. Consequently histidine showed no effect on the catalytic properties exhibited by the mutant enzyme. AMP on the other hand was still an effective inhibitor. *Cje*ATP-PRT Core even showed a higher sensitivity to AMP than the wild type when compared by their determined half maximal inhibitory concentration (IC_{50}) values. Taken together, these findings suggest that the truncation of the regulatory domain interferes with the catalytic properties of *Cje*ATP-PRT, largely without altering the active site ligand binding abilities. This compares well to the reported kinetic anal-

ysis of partially and totally truncated *Cgl*ATP-PRT mutants, which were found to possess a largely decreased specific activity and turn over in the crude cell extracts.[98]

The raw absorbency change of *Cje*ATP-PRT Core, compared to the *Cje*ATP-PRT wild type data, showed an early retention of the reaction rate before total turn-over was achieved. It was determined that this unexpected behaviour is due to a inhibitory effect of the reaction product PR-ATP. The (IC_{50}) value for PR-ATP was determined to be 6.8 $\pm$ 0.7 μM in the presence of 2 mM ATP (section 6.3.3), which shows the strong inhibitory potency of the product.

3.7 Ligand binding

3.7.1 Differential scanning fluorimetry

DSF experiments on the *Cje*ATP-PRT Core were performed as previously described for the wild type *Cje*ATP-PRT (section 6.1.37). For optimal comparison both enzymes were measured simultaneously on the same sample plate. The T_m of the *Cje*ATP-PRT Core mutant was determined to be 46.7 °C in the presence of 5 mM $MgCl_2$, which is 11 °C lower than the measured T_m of the wild type in the same condition. This finding shows that the removal of the ACT domain has a negative influence on the thermostability.

As shown for the wild type enzyme (Table 2.6), the presence of Mg^{2+} ions stabilises the enzyme, providing an increase of approximately 2.5 °C of the T_m value at 5 mM $MgCl_2$ concentration compared to the $MgCl_2$-free sample. This underlines the importance of Mg^{2+} for the stability and function of the enzyme. All other ligand samples therefore contained 5 mM Mg^{2+}.

The presence of a saturating ATP concentration did not increase the stability of the enzyme further. The resulting data (Figure 3.8) also shows conclusive evidence for the conservation of the AMP binding ability in the *Cje*ATP-PRT Core mutant through an increase of its thermostability by

4.5 °C. A stabilisation by histidine, as observed for the wild type enzyme, does not occur, neither in the presence nor in the absence of AMP. This is expected, as *Cje*ATP-PRT Core lacks a histidine binding site (Table 3.2).

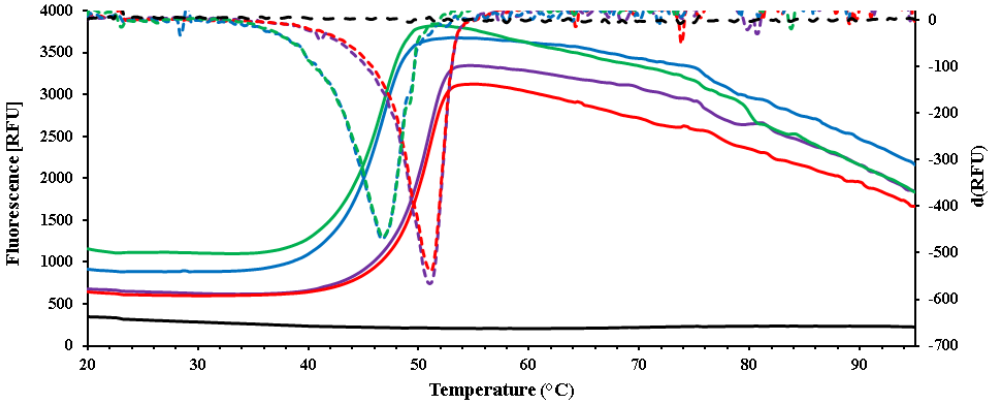


Figure 3.8: DSF analysis of *Cje*ATP-PRT Core. Melting curves (solid lines) of 0.1 mg mL $^{-1}$ enzyme samples were determined by DSF according to section 6.1.37 in the presence of 5 mM $MgCl_2$ and no ligand (green), 2 mM AMP (red), 1 mM histidine (blue) and a mix of 2 mM AMP and 1 mM histidine (purple). The first derivatives (dotted lines) show the inflection points of the melting curves, which represent the observed average T_m values.

Table 3.2: Thermostability of *Cje*ATP-PRT Core. Melting temperature values as determined by DSF. All samples contained 5 mM Mg^{2+}, unless stated otherwise. Samples were prepared and run according to section 6.1.37.

Additives	T_m
No ligand, no Mg^{2+}	44.3 ± 0.1 °C
No ligand, 0.5 mM Mg^{2+}	45.3 ± 0.1 °C
No ligand, 5 mM Mg^{2+}	46.7 ± 0.1 °C
2 mM ATP, 5 mM Mg^{2+}	46.6 ± 0.2 °C
2 mM AMP, 5 mM Mg^{2+}	51.2 ± 0.1 °C
1 mM histidine, 5 mM Mg^{2+}	46.7 ± 0.1 °C
2 mM AMP, 1 mM histidine, 5 mM Mg^{2+}	51.1 ± 0.1 °C

3.7.2 Differential scanning calorimetry

The collected DSF data was further supported by DSC experiments, which were conducted as described (section 6.1.38). The sample buffer contained 5 mM Mg^{2+}. T_m values were determined in the absence (52.5 ± 0.1 °C), and presence of 2 mM AMP (54.6 ± 0.1 °C), 1 mM histidine (52.4 ± 0.1 °C) and a mix of the latter two (55.0 ± 0.1 °C) (Figure 3.9). These melting temperatures were 4–6 °C higher than the values obtained from DSF, which is similar to the difference seen in the wild type experiments (section 2.7.2) and can be reasoned by the faster temperature ramping speed of the DSC experiments. The presence of histidine did not alter the T_m of *Cje*ATP-PRT Core, neither alone nor in combination with AMP, while 2 mM AMP increased the T_m by 2–2.5 °C.

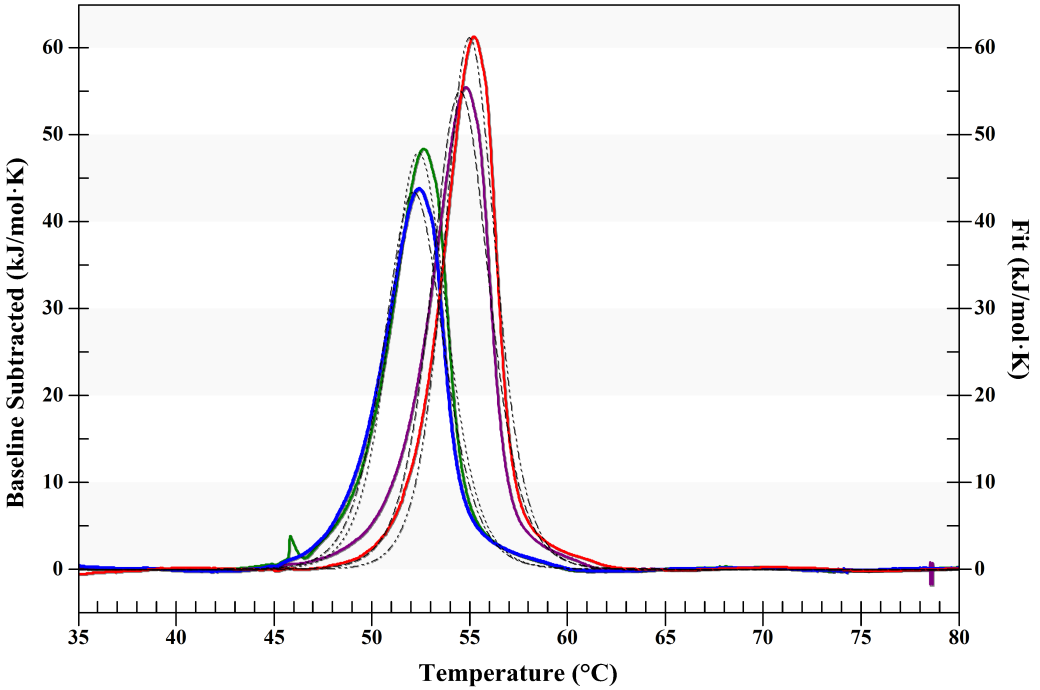


Figure 3.9: DSC analysis of *Cje*ATP-PRT Core. DSC measurements were performed according to section 6.1.38 with a $MgCl_2$ concentration of 5 mM. The resulting melting curves were baseline subtracted (solid lines) and fitted against a single melting event (dashed lines). Lines are coloured according to the ligand present in the experiment: none (green), histidine (blue), AMP (red), histidine and AMP (purple).

3.7.3 Isothermal titration calorimetry

Further confirmation for the binding of AMP to the *Cje*ATP-PRT Core mutant was obtained by ITC experiments. Using 2 mM AMP as titrant, a K_d value of 44 ± 2 μM was determined, which compares well to the wild type K_d value of 51 ± 4 μM. As for the wild type enzyme, the binding followed the expected single site model (Figure 3.10A) with a resulting stoichiometry of 0.99 ± 0.04. With no significant difference in the binding affinity of AMP to the macromolecule, it is evident that the active site is well suited for the accommodation of the inhibitor molecule AMP even in the absence of the regulatory domains and any cross-dimer contacts (α1-β2 loop "cap").

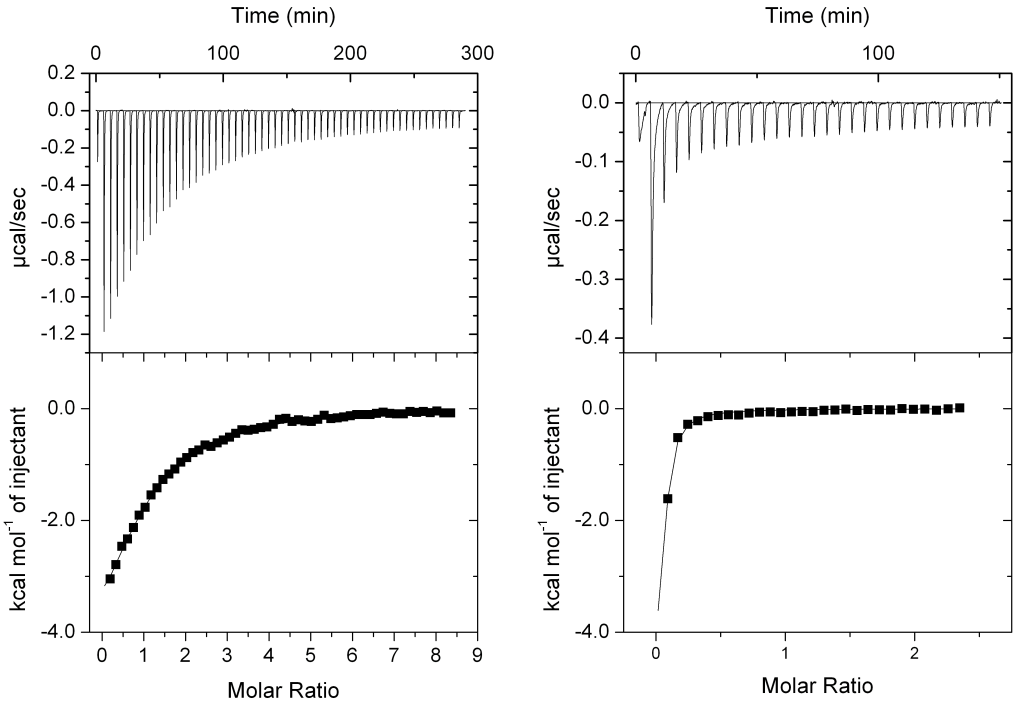


Figure 3.10: Nucleotide binding of *Cje*ATP-PRT Core. ITC binding curves of 2 mM AMP (A) and 2 mM ATP (B) against 100 μM *Cje*ATP-PRT Core were measured under the conditions described in section 6.1.39, subtracted by their corresponding heat of dilution data series and fitted using OriginLab® version 7.0.

ATP binding was also observed (Figure 3.10B) with *Cje*ATP-PRT Core. The binding was tighter than expected from the kinetic analysis (K_M(ATP)

= 117 ± 13 μM), which resulted in a very early saturation of the enzyme. Principally it can be stated, that *Cje*ATP-PRT Core is able to bind the substrate molecule ATP. The K_d value obtained from the data fit was 10.0 ± 4.3 μM, but should be disregarded as there were not enough suitable data points available. Using a lower ATP concentration, it should be possible to acquire a more accurate K_d value. Unfortunately this has not be achieved due to time constraints and the large quantity of enzyme required for this measurement.

3.8 Crystallographic analysis

3.8.1 Crystallisation

Crystallisation procedures were performed as described in section 6.3.5. During initial screening, the condition A12 of the PACT *premier*™ HT-96 was identified as yielding protein crystals of undesirable properties. The condition solution was rich in heavy, brown precipitate and the crystals formed were of dark colour, small and grew as multiples. Due to the lack of other suitable conditions, attempts were made to optimise this condition. By reducing the precipitant concentration from 20 % to 7–10 % w/v and increasing the initial protein concentration in the drop to $\geq$ 10 mg mL $^{-1}$, large single clear crystals were grown over night from heavy precipitate, that formed immediately after the preparation of the drops. The shape of the crystals was perfectly octahedral (Figure 3.11), unless crystal growth was stopped by contact with a surface. Crystals grew consistently from either fresh or frozen protein stocks and were suitable for ligand soaking. ATP, PRPP and PR-ATP were soaked into the crystals by the addition of concentrated stock solutions directly to the drop. Generally the crystals did not break up or dissolve during this treatment and were looped, cryoprotected and flash frozen 30–60 minutes afterwards.

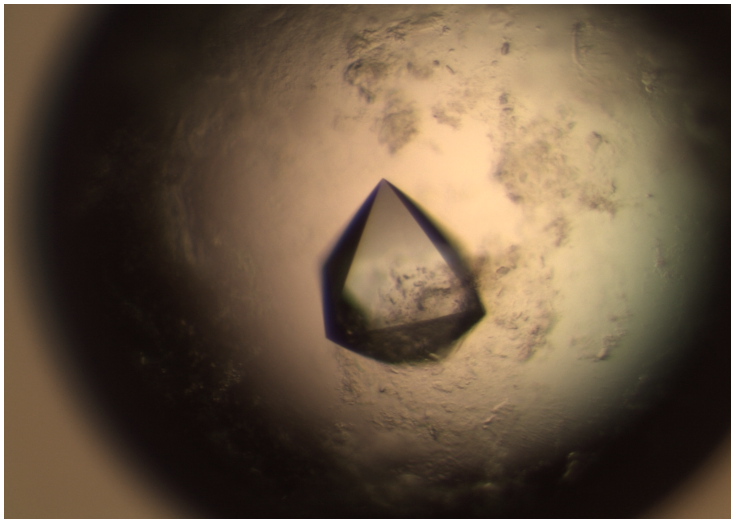


Figure 3.11: Octahedral crystal of *Cje*ATP-PRT Core. Photograph taken of a drop containing a single *Cje*ATP-PRT Core crystal amidst left-over precipitate. The octahedral crystal was obtained from the adjusted PACT *premier*™ HT-96 screen condition A12: 0.1 M Sodium Acetate, 0.01 M $ZnCl_2$, 7 % w/v PEG 6000, pH 5.0. The initial protein concentration in the drop was 11 mg mL $^{-1}$.

It was not possible to combine AMP with the crystal condition, by either ligand soaking or as a co-crystallisation agent. The condition solution transformed into a gel upon addition of AMP, which was not able to diffuse into the crystals. The cause of the gel formation was identified to be the Zn^{2+} contained in the crystallisation condition, but Zn^{2+} was also found to be crucial for the crystallisation process. Because of the excellent diffraction quality of the crystals described above and the initial difficulties to find appropriate crystallisation conditions, no further screening to introduce AMP was attempted.

3.8.2 Crystal structure

The first X-ray dataset of *Cje*ATP-PRT Core was collected on an apo crystal with diffraction spots distributed out to 1.6 Å. The data was later cut back to 1.9 Å and was integrated in space group $P2_12_12_1$. A solution was found by molecular replacement using a search model created from a single chain of

Table 3.3: Crystal parameters, data collection, and refinement statistics for *Cje*ATP-PRT Core.

	*Cje*ATP-PRT Core
Data collection	
Crystal system	trigonal
Space group	$P2_12_12_1$
Unit cell parameters	
a, b, c [Å]	65.70, 79.93, 91.01
α, β, γ [°]	90.00, 90.00, 90.00
Resolution range [Å]	40.00–1.90 (1.94–1.90)
Measurements	276691
Unique reflections	38485
Completeness	100.00 (99.6)
Redundancy	7.2 (6.6)
$I/\sigma(I)$	22.9 (4.8)
R_{merge}	0.050 (0.321)
$CC_{1/2}$	0.999 (0.956)
Wilson B value [Å^2]	23.3
Matthews coefficient	2.36
Refinement	
Resolution [Å]	40.00–1.90
R_{cryst}	0.193
R_{free}	0.243
Chain length	226
Observed residues	220, 215
Water molecules	295
Others	5 Acetate, 3 Zn^{2+}
Mean B [Å^2]	
Protein	28.08
Other	36.39
Water	31.50
RMSD	
Bond lengths [Å]	0.013
Bond angles [°]	1.582
Chiral volumes [Å^3]	0.088
Ramachandran	
Preferred %	98.62
Allowed %	1.38
Outliers %	0.00
PDB entry	4YRX

the wild type structure 4YB5 that had the residue range 225–299 removed. Two chains were found in the unit cell sharing one interface identical to the described dimer interface of the wild type *Cje*ATP-PRT. The entire helix $\alpha 8$ of chain B showed significant variability compared to the rest of the structure, which led to the assignment of a separate TLS group to it during refinement. Later ligand bound structures were solved from the *Cje*ATP-PRT dimer (4YRX), again by molecular replacement.

3.8.3 Fold and arrangement

The descriptions and comparisons made in the following section largely correspond to the *Cje*ATP-PRT Core apo structure (4YRX) but all ligand bound structures contain very similar features (Table 3.4). Although the two chains resolved in the unit cell show the identical secondary structure arrangement of domain I and II as described for the wild type (section 2.9.3), they present two main differences with an overall RMSD between chains A and B of 0.38 Å. The more prominent of these is the shorter C-terminus of chain B when compared to chain A. No electron density was observed for residues 220–225 in chain B. Therefore these residues were removed from the model.

Chain A shows an alternative conformation for the $\beta 9$-$\alpha 7$ loop (the PRPP binding loop), which was found stretched out, lacking the typical narrow turn seen in chain B, and all other ATP-PRT structures. The subsequent $\alpha 7$ helix is also effected by this change, being largely unwound with only one helical turn remaining (residues 178–181). This conformation was only observed in chain A of 4YRX, not in any of the ligand bound *Cje*ATP-PRT Core structures, where the loop is performing its binding function. It is possible that this alternative conformation represents a "resting" position of this important binding loop. A significant feature of the structure was the coordination of several Zn^{2+} ions at the unit cell interfaces, interacting mainly with residues H33 and H35 as well as E78 and E81, but also with residues of the $\beta 9$-$\alpha 7$ loop of chain A, which could in part be the reason for its difference in conformation.

Table 3.4: Comparison of the *Cje*ATP-PRT Core structures. Comparison of the dimeric subunit arrangements comprising residues 4–225 of all listed *Cje*ATP-PRT structures is given as RMSD value derived from their superimposition using PyMol.[99] PDB identifiers are used for the structures of *Cje*ATP-PRT Core (4YRX), histidine bound *Cje*ATP-PRT wild type (4YB5), AMP and histidine bound *Cje*ATP-PRT wild type (4YB6) and ATP bound *Cje*ATP-PRT wild type (4YB7).

Crystal structure	4YRX	4YB5	4YB6	4YB7
4YRX	-	1.61 Å	2.17 Å	3.78 Å
PR-ATP bound *Cje*ATP-PRT Core	0.47 Å	1.45 Å	2.14 Å	3.63 Å
ATP bound *Cje*ATP-PRT Core	0.47 Å	1.54 Å	2.19 Å	3.70 Å
PRPP bound *Cje*ATP-PRT Core	0.41 Å	1.55 Å	2.22 Å	3.80 Å

3.8.4 Comparison to the *Cje*ATP-PRT wild type

When compared to the wild type *Cje*ATP-PRT structures, several distinguishing features of the Core mutant are evident. Logically the resolved chains entirely lack the regulatory ACT domain. Most importantly though, the presented dimer interface closely resembles the dimer interfaces of the histidine bound wild type structures (4YB5 and 4YB6). Superimposition of the *Cje*ATP-PRT Core crystal structures with the dimeric catalytic core portions of the wild type structures (Table 3.4) shows the highest similarities exist between all Core mutant structures the histidine bound wild type structure (4YB5). In these structures the same close distance between the C-terminal ends of α3 (3.4 Å between the peptide oxygen atoms of S86 and S86') and large distances between α8 helices and the α1-β2 loops (12.3 $\pm$ 0.4 Å between peptide oxygen of A179 and γ oxygen of S38') are found, which stand in stark contrast to the measurements made on the ATP bound wild type structure (4YB7) (14.0 $\pm$ 0.7 Å between the peptide oxygens of S86 and S86', and 5.2 $\pm$ 0.7 Å between peptide oxygen of A179 and γ oxygen of S38'). This inverse relationship shows that the *Cje*ATP-PRT Core mutant

exists in a similar conformation to the inhibited dimeric units of the wild type (Figure 3.12). This was observed for the unliganded as well as for all ligand bound crystal structures of *Cje*ATP-PRT Core.

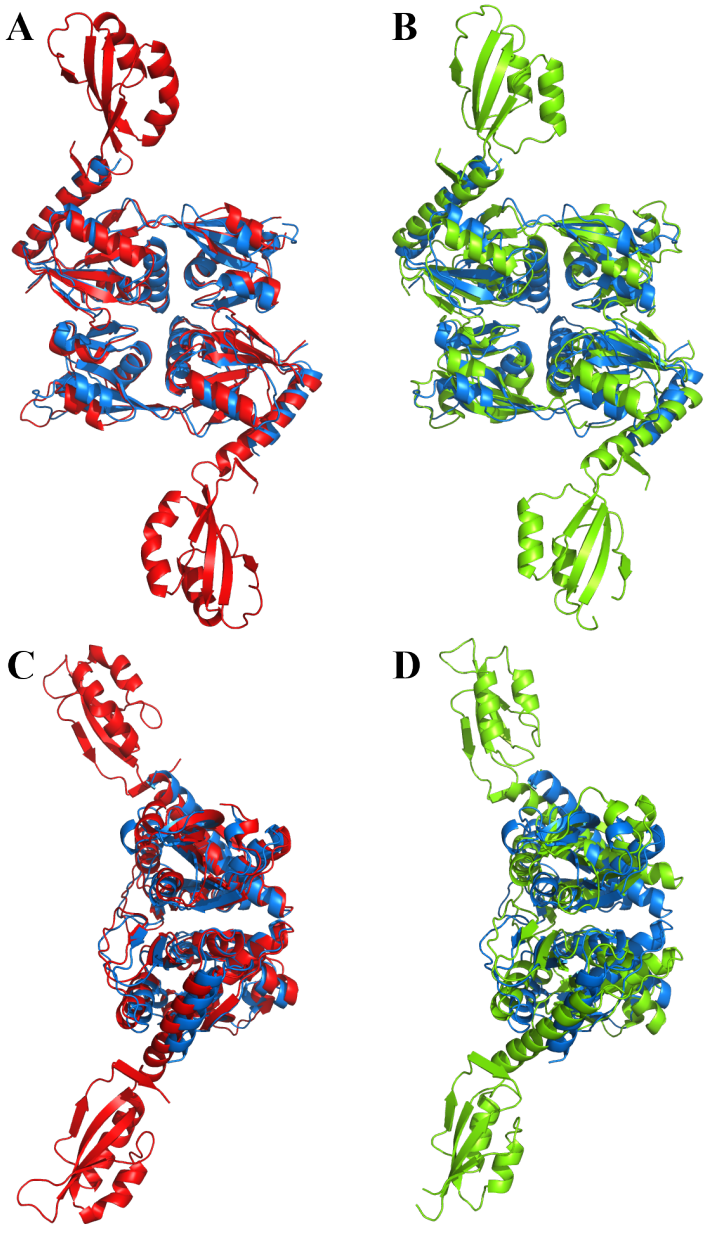


Figure 3.12: Structural superimposition of *Cje*ATP-PRT Core and wild type enzymes. Front (A + B) and side view (C + D) on the dimeric assemblies of 4YRX (blue), 4YB5 (red) and 4YB7 (green) in their cartoon representation. Overlays were performed using PyMol.[99]

The structure also supports the existence of the mutant as a dimer, as the dimers of neighbouring unit cells fail to arrange in a composition resembling the *Cje*ATP-PRT wild type hexamer.

3.8.5 Active site ligands

The ligand soaked crystals of *Cje*ATP-PRT Core resulted in data sets of comparable resolution to the apo structure (Table 3.5), featuring well resolved ligand density for the substrates ATP and PRPP, but especially for the product PR-ATP (Figure 3.15A). A summary of their parameters and refinement statistics can be found in Table 3.5. The position of the ligands and the involved binding interactions in the active site are described in detail in this section.

The substrate ATP is bound in a position nearly identical to its binding mode in the wild type enzyme (Figure 3.13). Both orientations of the adenine that had been observed in 4YB7 were also found in *Cje*ATP-PRT Core, with a phosphate ion bound in the loop next to the ATP in the catalytically irrelevant position. Consequently all interactions of ATP described in section 2.9.4 with active site residues are also present in *Cje*ATP-PRT Core. Residues Q12, R16 and R54 bind to the triphosphate, while N75 forms hydrogen bonds to the ribose, and C104, L170 and R16 form hydrophobic interactions with the adenine ring system. The metal in complex with ATP differs. Although the binding site between D55 and D56 is equivalent, the high electron density of the coordinated metal in *Cje*ATP-PRT Core was attributed to the binding of Zn^{2+} from the crystallisation condition. The zinc ion substitutes the magnesium ion, found in the wild type enzyme, and adopts an octahedral coordination geometry uncommon for Zn^{2+}. These results strongly support the binding experiments performed by ITC and confirm that ATP binding is not adversely affected by the removal of the regulatory domain.

Table 3.5: Preliminary crystal parameters, data collection, and refinement statistics for *Cje*ATP-PRT Core ligand bound structures.

	ATP	PRPP	PR-ATP
		Data collection	
Crystal system	trigonal	trigonal	trigonal
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters			
a, b, c [Å]	66.17, 79.92, 92.25	66.00, 79.69, 91.12	67.46, 80.46, 92.73
α, β, γ [°]	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution range [Å]	46.12–2.00(2.05–2.00)	45.56–2.10(2.16-2.10)	44.55–1.90(1.94–1.90)
Measurements	264717	422010	439398
Unique reflections	33714	28897	40502
Completeness	99.8 (100.00)	99.9 (98.5)	99.8 (97.1)
Redundancy	7.9 (8.1)	14.6 (14.4)	10.8 (10.5)
$I/\sigma(I)$	13.6 (3.3)	27.3 (3.7)	23.9 (3.2)
R_{merge}	0.082 (0.549)	0.096 (0.803)	0.067 (0.910)
$CC_{1/2}$	0.998 (0.948)	0.999 (0.896)	1.000 (0.769)
Wilson *B* value [Å^2]	31.4	28.6	24.8
Matthews coefficient	2.41	2.37	2.49
		Refinement	
Resolution [Å]	40.00–2.00	40.00–2.10	40.00–1.90
R_{cryst}	0.208	0.209	0.196
R_{free}	0.251	0.263	0.229
Chain length	226	226	226
Observed residues	222, 216	222, 218	222, 222
Water molecules	148	207	194
Others	2 ATP, 1 phosphate, 1 ribose, 3 Zn^{2+}, 3 acetate	2 PRPP, 3 acetate, 6 Zn^{2+}, 2 Mg^{2+}, 2 1,2-ethylendiol	2 PR-ATP, 6 Zn^{2+}
Mean *B* [Å^2]			
Protein	40.88	32.80	34.15
Water	39.35	34.00	32.12
Other	49.46	40.88	48.87
Ligand	52.55	54.78	31.20
RMSD			
Bond lengths [Å]	0.013	0.012	0.008
Bond angles [°]	1.710	1.504	1.430
Chiral volumes [Å^3]	0.090	0.082	0.078
Ramachandran			
Preferred %	98.38	97.13	97.93
Allowed %	1.39	2.87	1.84
Outliers %	0.23	0.00	0.23

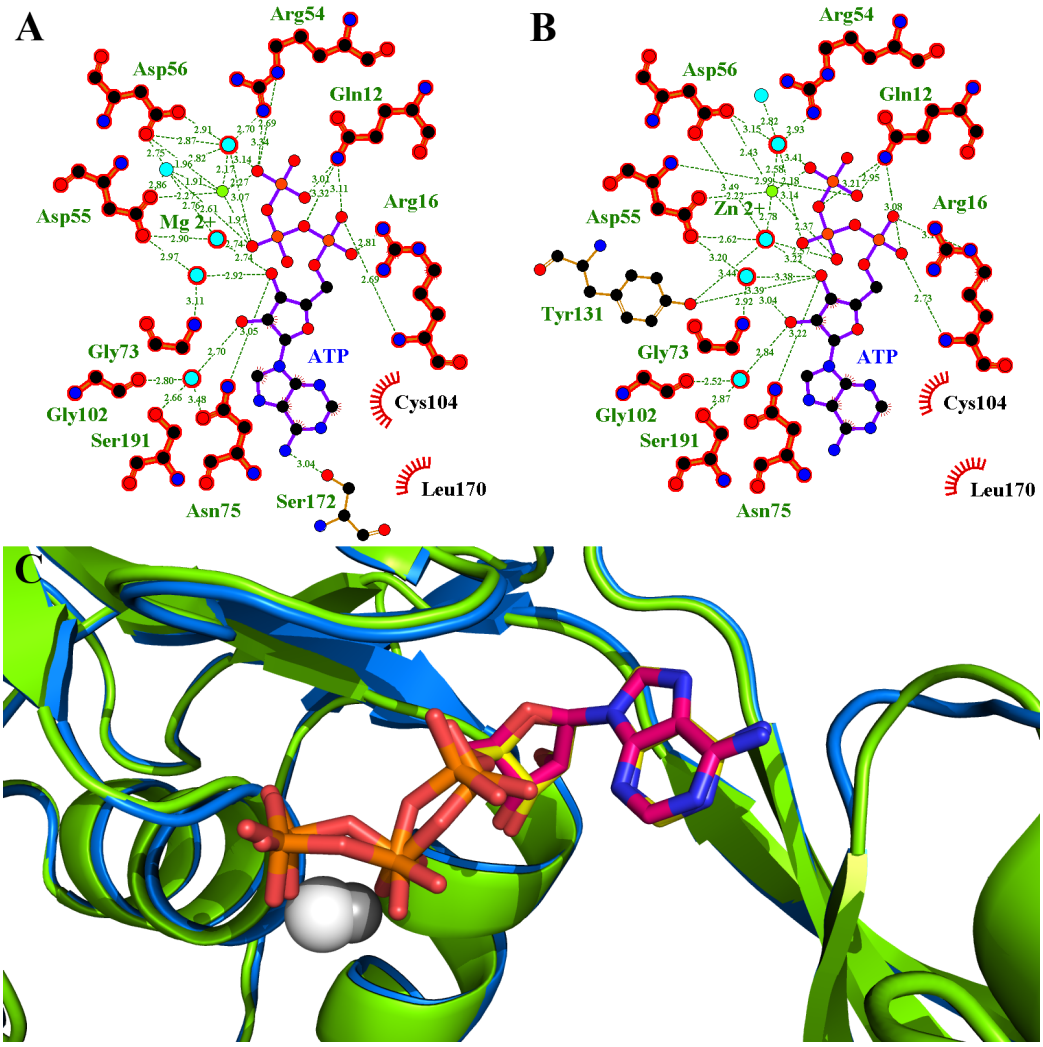


Figure 3.13: ATP binding in *Cje*ATP-PRT Core. The binding mode of the ATP molecules found in 4YB7 (A) and ATP bound *Cje*ATP-PRT Core (B) are compared by A/B: Planar plots of all ATP interactions in both structures created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with the real space distances between atoms. Hydrophobic interactions are displayed as short radial red lines. Conserved interactions are highlighted in red. Water molecules are displayed as cyan spots. Metal ions are green. C: Superimposition of the three-dimensional structures of 4YB7 (backbone green, ATP pink) and ATP bound *Cje*ATP-PRT Core (backbone blue, ATP yellow) in cartoon and stick representations. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), magnesium (white), zinc (grey).

The substrate PRPP has not been observed in any ATP-PRT long form crystal structure reported to date, but correlations can be made to the short form structure of *Lla*ATP-PRT, as the active sites are similar in long and short forms. The binding mode of PRPP observed in the active site of *Cje*ATP-PRT Core (Figure 3.14) was different from the expectations based on previous findings.[36] The 5'-phosphate binds in the PRPP binding loop, as expected, but the pyrophosphate bridges over to the metal coordination site on domain I, instead of interacting with S154 and E156 of domain II.

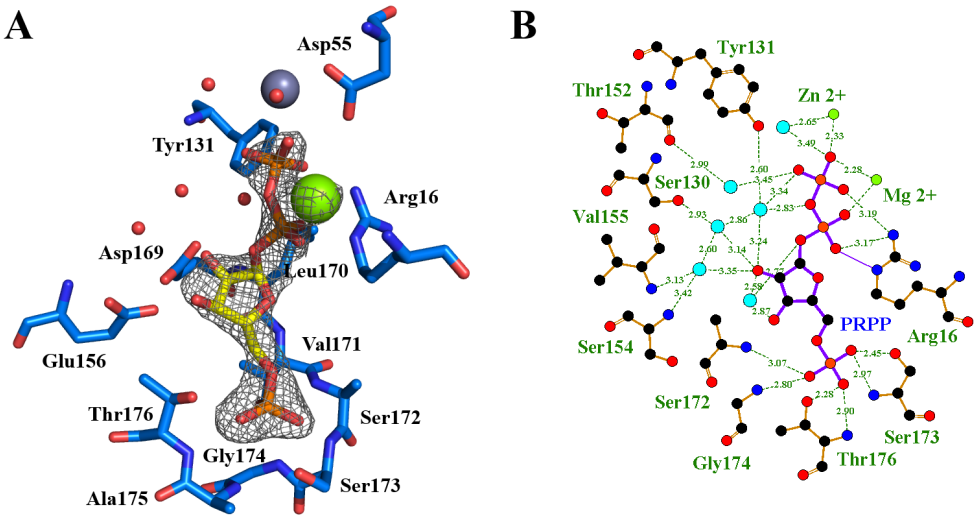


Figure 3.14: Observed PRPP binding mode. A: PRPP (yellow) surrounded by electron density (Fo-Fc map - grey mesh). Active site residues of *Cje*ATP-PRT Core (blue) are shown as sticks. Zn^{2+} (grey) and Mg^{2+} (green) ions are displayed as spheres. B: Planar plot of all PRPP interactions created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with the real space distances between atoms. Hydrophobic interactions are displayed as short radial red lines. Water molecules are displayed as cyan spots. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

In the observed orientation, PRPP would cause steric clashes with ATP bound in the conformation described above. Therefore this conformation may not be catalytically relevant. Most of the interactions seen for the ribose-5'-phosphate portion of PRPP are identical to the interactions described for AMP in the wild type (section 2.9.4), mainly formed with the backbone and short side chains of residues 169–176 (the PRPP binding loop). The py-

rophosphate group forms interactions with Y131 and R16 and the bound metal ions. R16 bends out of its ATP binding conformation to accommodate the pyrophosphate group of PRPP, which reflects the importance of positively charged side chains for the binding of the phosphate-rich ATP-PRT substrates.

As for the other ligand bound crystal structures, the PRPP bound *Cje*ATP-PRT Core contains zinc in the metal coordination site (D55, D56), but an additional magnesium ion is observed binding between two oxygen atoms of the PRPP α and β-phosphate. The presence of magnesium in addition to zinc was unexpected, but is likely due to the use of Mg^{2+} saturated PRPP solution for the crystal soak.

This is the first report of the large product molecule PR-ATP as a complete entity in an ATP-PRT active site. The information about the PR-ATP binding mode gathered from this structure is comparable to the findings of the substrate binding modes. PR-ATP is accommodated in both parts of the active site. It can be described roughly as "zig-zag"-shaped, with the adenine in the centre and the ribose monophosphate and the ribose triphosphate on either end facing in opposite directions (Figure 3.15).

This conformation closely resembles the superimposition of ATP and AMP binding modes found in the *Cje*ATP-PRT wild type structures. The N_1 linked phosphoribosyl group binds in the same position as seen for PRPP and AMP, making contacts to residues of the PRPP binding loop. The adenine is held between R16 and L170 in a position equivalent to the ones seen for ATP and AMP. The ribose triphosphate part of the molecule sits in a position again identical to that of ATP.

The only unoccupied area in each active site is situated next to the phosphoribosyl moiety on domain II, formed by the β8-α6 loop and the conserved residue E156 (of helix α8). This most likely represents the binding site for PP, either free or as part of PRPP. The metal coordination site is also identical to the other ligand bound structures of *Cje*ATP-PRT Core and contains a zinc ion.

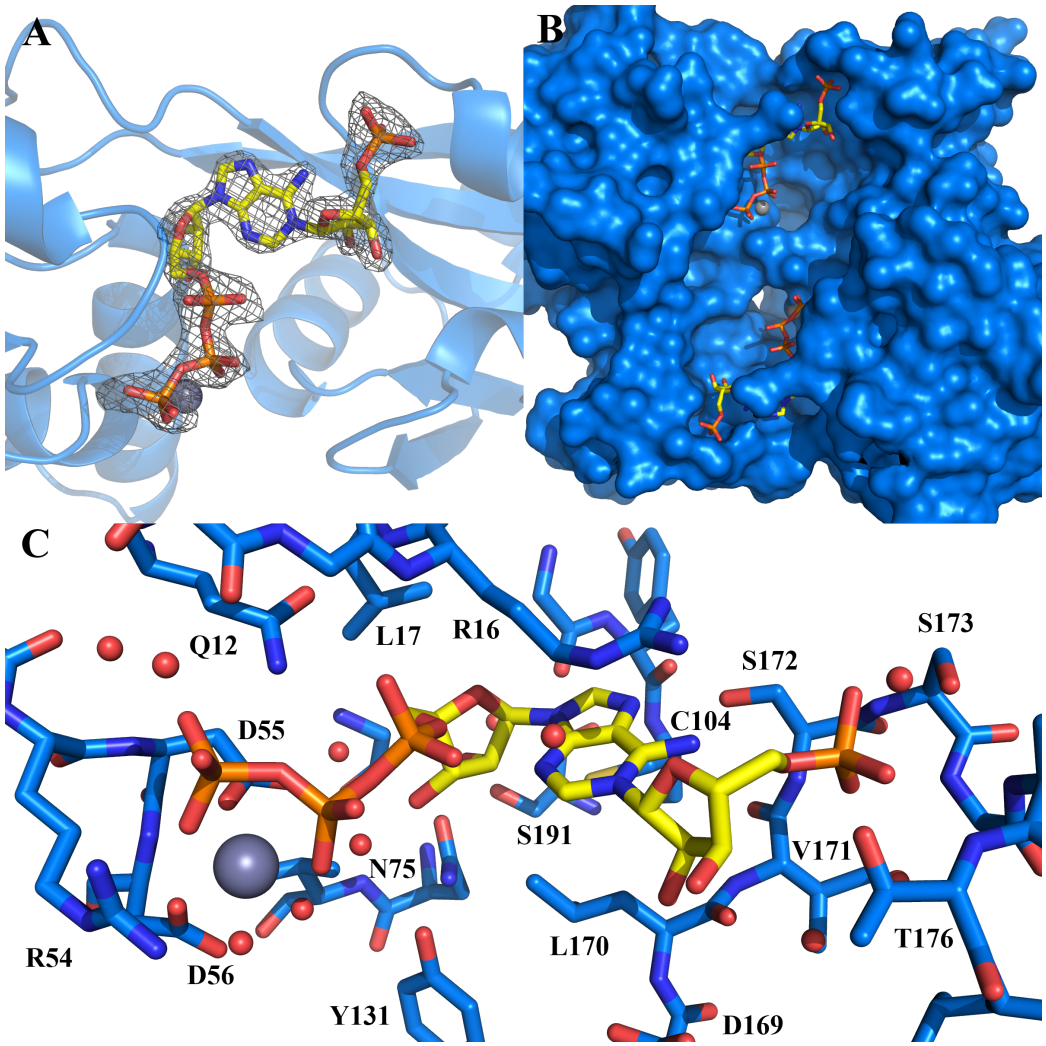


Figure 3.15: Phosphoribosyl-ATP in the active site of *Cje*ATP-PRT Core. A: PR-ATP (bright yellow) surrounded by the observed electron density (Fo-Fc map - grey mesh). The *Cje*ATP-PRT Core (blue) is shown in cartoon representation and the Zn^{2+} ion as a sphere (grey). B: Surface display of the dimer with the PR-ATP molecules bound in the active site cavities. C: PR-ATP and Zn^{2+} binding mode in the active site. Residues and water molecules in close distance are displayed as sticks and spheres respectively. Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange), sulphur (yellow), zinc (grey).

PR-ATP showed a strong, well defined electron density with the average B factor of 31.2 Å^2 well matched to the surrounding residues (25-35 Å^2). This is indicative of a high affinity to the active site and low conformational flexibility. This finding strongly supports the determined inhibitory properties of PR-ATP for *Cje*ATP-PRT Core.

3.9 The *Cje*ATP-PRT ACT domain

3.9.1 Cloning and purification

Alongside the *Cje*ATP-PRT Core the second portion of the full length protein chain, the regulatory domain, was cloned as a separate protein (Figure 3.16A) denoted *Cje*ATP-PRT ACT. The cloning followed the procedure described in section 6.3.1, with two different forward primers, to generate a tagged construct using vectors pDEST™15 and pDEST™17, and an untagged construct in pDEST™14. As for the Core mutant, a desirable level of solubility of the overexpressed protein was only achieved using the GST-tagged construct (pDEST™15), which was consequently used for all purifications of *Cje*ATP-PRT ACT.

The purification protocol for the *Cje*ATP-PRT ACT protein followed a sequence of chromatography steps. An initial GST affinity chromatography step isolated the protein from the clarified supernatant. It was followed by TEV protease incubation (1 h at 37°) and a second GST affinity chromatography to remove the free GST tag. After these steps the protein preparation still contained TEV protease, which co-eluted with the *Cje*ATP-PRT ACT from a SEC run. Therefore the protein preparation underwent further purification by IMAC to remove the TEV protease before the final SEC step (Figure 3.16B).

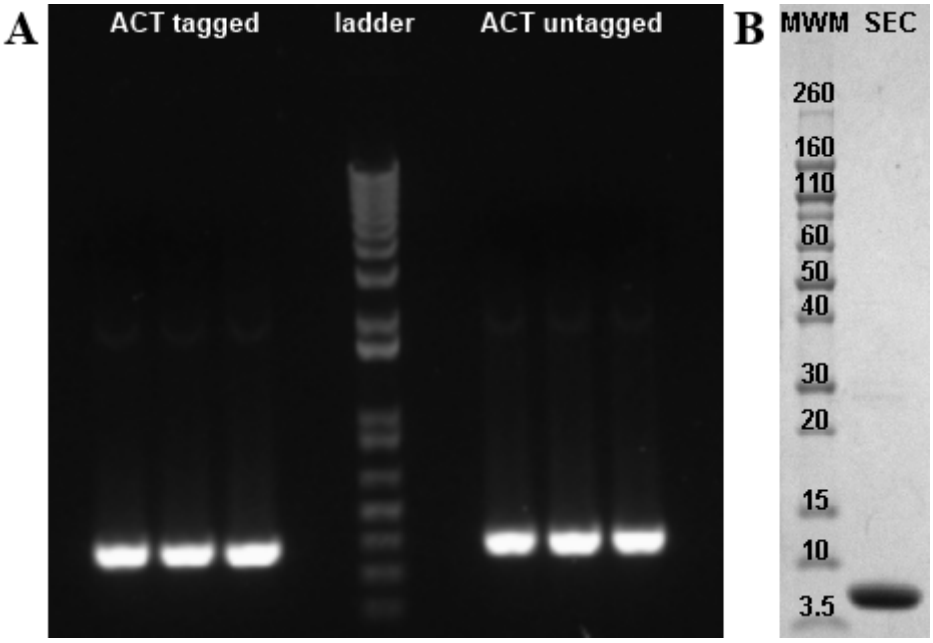


Figure 3.16: Preparation of *Cje*ATP-PRT ACT. A: Agarose gel showing the amplified genes of *Cje*ATP-PRT ACT (tagged and un-tagged) from three selected clones each using the colony PCR set up described in section 6.1.17. The gel was run according to section 6.1.13. B: Protein purity was assessed by SDS-PAGE, ran according to section 6.1.31 using a Bolt® 10% Bis-Tris Gel and MES buffer.

3.9.2 Characterisation

It was noted during its purification that the purified *Cje*ATP-PRT ACT species co-elutes with TEV protease (27 kDa) from a size exclusion column. As the theoretical molecular mass of *Cje*ATP-PRT ACT is 8.7 kDa. This was taken as a first indication that it may exist as a trimer, analogous to the regulatory trimer arrangement of the full length *Cje*ATP-PRT enzyme. To confirm the oligomeric state, SLS was performed, resulting in a calculated mass of 23.4 kDa, which is closer to a trimer (26.1 kDa) than to a dimer (17.4 kDa). Consequently it was assumed that the protein exists in a trimeric form in solution.

*Cje*ATP-PRT ACT was tested for its thermostability using DSC. A single run containing 0.35 mg mL $^{-1}$ protein in the cell was performed following the description in section 6.138. The resulting curve showed a very

late melting event for the protein sample around 94 °C (Figure 3.17), which was confirmed by later repeats. This unexpectedly high melting temperature was interpreted as a measure of tight internal interactions of the ACT domain trimer, which lead to a very stable conformation.

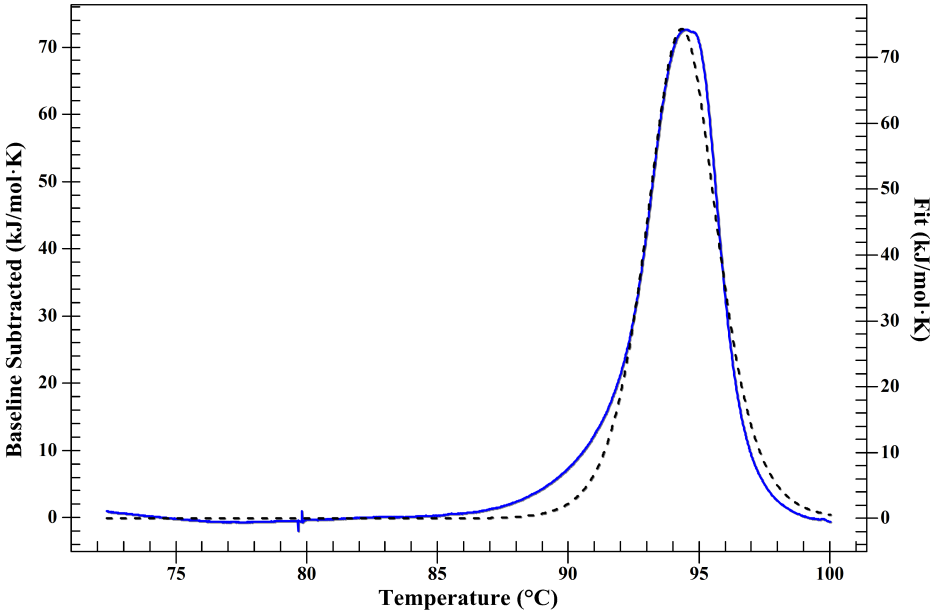


Figure 3.17: Thermostability of *Cje*ATP-PRT ACT. DSC measurements were performed according to section 6.1.38 with a $MgCl_2$ concentration of 5 mM. The resulting melting curve was baseline subtracted (solid line) and fitted against a single melting event (dashed line).

To assess the function of the pure *Cje*ATP-PRT ACT preparation it was used in a kinetic assay with *Cje*ATP-PRT Core (section 6.3.3) at *Cje*ATP-PRT Core to *Cje*ATP-PRT ACT ratios of 1:1, 1:10, and 1:20. The outcome was very clear. The ACT domain did not display any measurable activity itself, nor did it enhance the activity of the *Cje*ATP-PRT Core mutant at any of the tested ratios. It also did not impose a sensitivity towards histidine on the Core mutant (Figure 3.18B).

To test if the purified protein was functional as a histidine binding module, an ITC experiment was performed. Under the given conditions (section 6.3.4) binding of histidine to *Cje*ATP-PRT ACT was not observed (Figure 3.18A). Together the kinetic and ITC results indicate the protein

does not appear to have retained the function of the covalently linked ACT domain in the *Cje*ATP-PRT wild type enzyme.

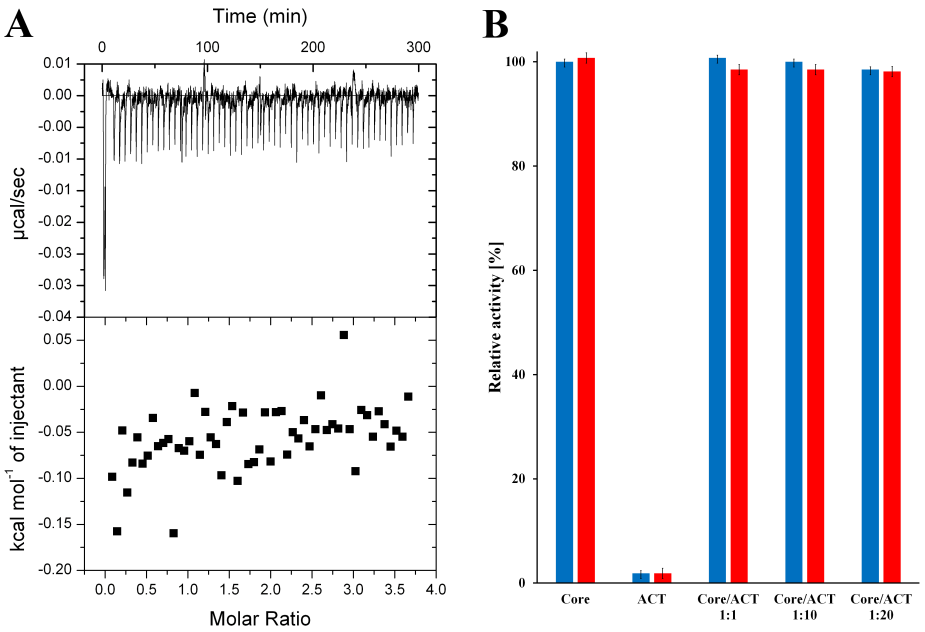


Figure 3.18: *Cje*ATP-PRT ACT functionality. A: ITC binding curve of 1.5 mM histidine against 57 μM protein was measured under the conditions described in section 6.1.39, subtracted by the corresponding heat of dilution data series and fitted using OriginLab® version 7.0. B: Kinetic data acquired for *Cje*ATP-PRT Core and ACT under the assay conditions described in section 6.3.3 in the absence (blue) and presence (red) of 0.5 mM histidine. Data was normalised against the *Cje*ATP-PRT Core activity without additives.

Circular dichroism (CD) can be used to assess the secondary structure composition of a pure protein sample (section 6.1.36). The inconclusive results for the *Cje*ATP-PRT ACT protein led to comparative analysis by CD. The resulting ellipticity curve showed that the isolated protein species is at least partially unfolded, displaying a prominent peak in the CD spectrum around 204 nm (Figure 3.19). Peaks at 210 nm (α helix) and 220 nm (β strand) are only weakly pronounced, compared to the theoretical spectrum[100] for the *Cje*ATP-PRT ACT trimer, which is consistent with a loss of those secondary structure elements. Given its defined oligomeric state and high melting temperature, this result was surprising.

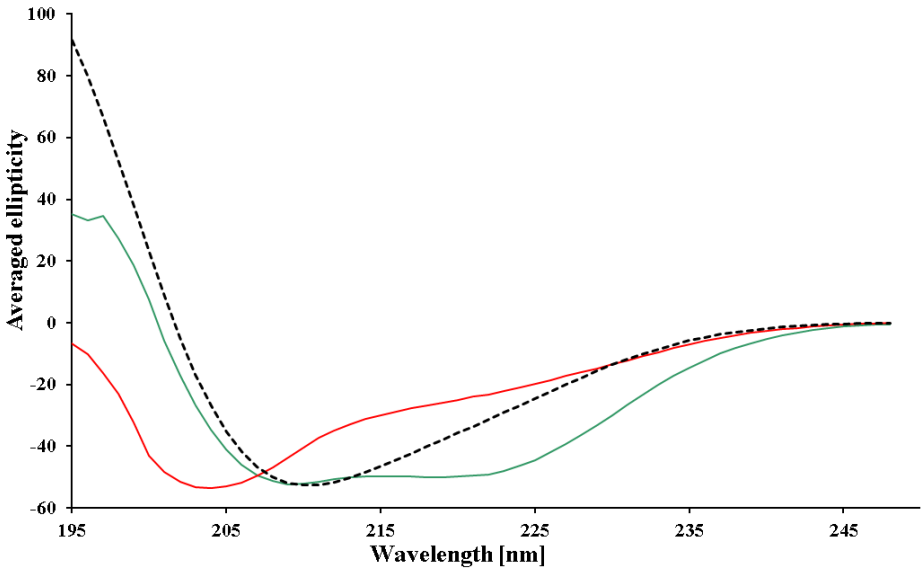


Figure 3.19: Folding of *Cje*ATP-PRT ACT. Five point averaged ellipticity was generated from the raw CD measurements according to section 6.1.36. The curves for *Cje*ATP-PRT wild type (green) and *Cje*ATP-PRT ACT (red) are shown in comparison. A predicted spectrum (black dashed) was generated from the PDB coordinates of an isolated ACT trimeric unit using DichroCalc.[100] Typically CD peaks are associated with: β strand (220 nm), α helix (210 nm) and random coil (200 nm).

3.9.3 Crystallisation

Crystallisation trials with *Cje*ATP-PRT ACT resulted in a couple of initial hits (section 6.3.5). All of these conditions were up-scaled and fine-screened but failed to produce any crystals except JCSG*plus premier*™ HT-96 B9. Crystals grown from this condition were of very poor diffraction quality, resulting in weak diffraction patterns with resolutions $\geq$ 10 Å.

A single data set was collected from a crystal grown directly in the JCSG*pluspremier*™ HT-96 screen in condition E1 (Figure 3.20). Diffraction spots were visible out to approximately 4 Å, but the overall data quality was too poor to resolve the structure by molecular replacement. If the protein is in fact partially or completely unfolded, a solution by molecular replacement with the *Cje*ATP-PRT ACT domain as model is very unlikely.

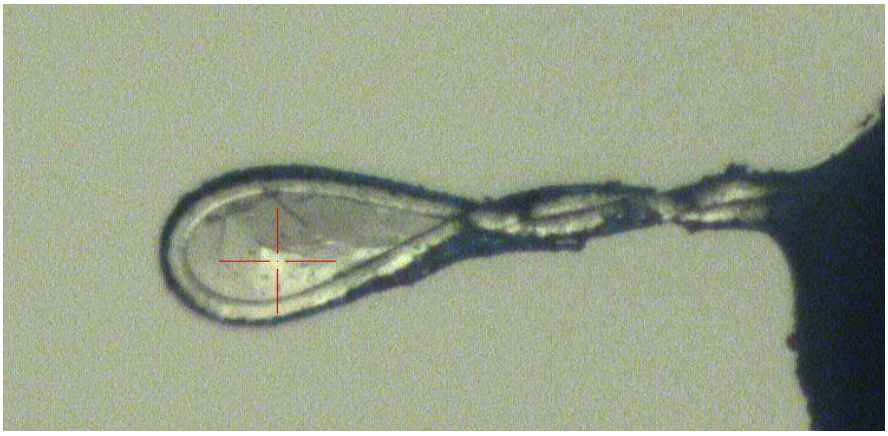


Figure 3.20: *Cje*ATP-PRT ACT crystal. Crystal image of *Cje*ATP-PRT ACT crystal from JCSG*plus premier*™ HT-96 screen condition E1 taken during diffraction collection on the Australian Synchrotron MX1 beam line.

3.10 Discussion

All results presented in this chapter are consistent with the formation of a dimeric, functional, truncated species of the *Cje*ATP-PRT enzyme by removing the regulatory ACT domain. The *Cje*ATP-PRT Core mutant shows low ATP-PRT activity and active site ligand interactions, but lacks the sensitivity to histidine, along with its entire regulatory apparatus. The presented crystal structures contain high quality information about the ligand binding properties of the *Cje*ATP-PRT active site. This includes the presence of the product molecule PR-ATP with high detail and resolution. Unfortunately under the conditions used, an AMP bound crystal structure could not be obtained, but all other results clearly show that AMP is still an effective inhibitor to the enzyme.

Zinc has been found to replace the active site magnesium ion in all ligand bound *Cje*ATP-PRT Core crystal structures. As part of the crystallisation condition, zinc is present during crystallisation in a 2:1 ratio with magnesium that is present in all *Cje*ATP-PRT buffers. The high affinity of zinc to the metal site can therefore not be attributed to its excess alone. Activity measurements with zinc in the absence of magnesium were not successful

due to the immediate formation of precipitate of both *Cje*ATP-PRT Core and wild type enzyme. The same protein-destabilising effect of zinc was seen during crystallisation. The fortunate exception here was that *Cje*ATP-PRT Core is able to refold and form high quality crystals from the precipitate. This suggests that zinc does not have any biological function in *Cje*ATP-PRT. Nevertheless zinc can substitute for magnesium in the active sites of *Cje*ATP-PRT Core crystals, adopting the octahedral coordination of magnesium in the presence of ligand molecules. It is likely that this substitution and the involvement of Zn^{2+} in the crystal packing provides stability to the complex and allows for the high resolution of the structures and the observation of the bound ligands.

PRPP was found to bind the active site in an unexpected way.[36] This is likely to be caused by the presence of zinc in the crystallisation condition, which attracts the negatively charged pyrophosphate group of the bound PRPP molecule. Similarly, a magnesium ion in the metal coordination site (D55 and D56) could cause the observed conformation. The additional magnesium ion bound to the α and β-phosphate supports this assumption. The way in which PRPP binds obscures the binding site for the ATP triphosphate (Figure 3.21A), as both ligands individually coordinate to the bound metal. This binding mode could represent the "safe keeping" of PRPP by the enzyme, discussed in chapter 2 and explain the observed binding of PRPP to the wild type enzyme in the absence of ATP. Upon the entrance of an ATP molecule, the pyrophosphate group could be removed from the metal by the electrostatic repulsion of the ATP triphosphate and shift PRPP into the expected position for catalysis (Figure 3.21B).

The *Cje*ATP-PRT Core preparation showed significant levels of activity under the used assay conditions, but very high protein concentrations were required. This is reflected in its low turn over rate of < 0.03 s^{-1}. Recent studies on *Cgl*ATP-PRT have shown that the truncation of the regulatory domain in the native enzyme leads to a comparable decrease in the catalytic turnover rates of the resulting mutants,[98] consistent with the results obtained for *Cje*ATP-PRT Core.

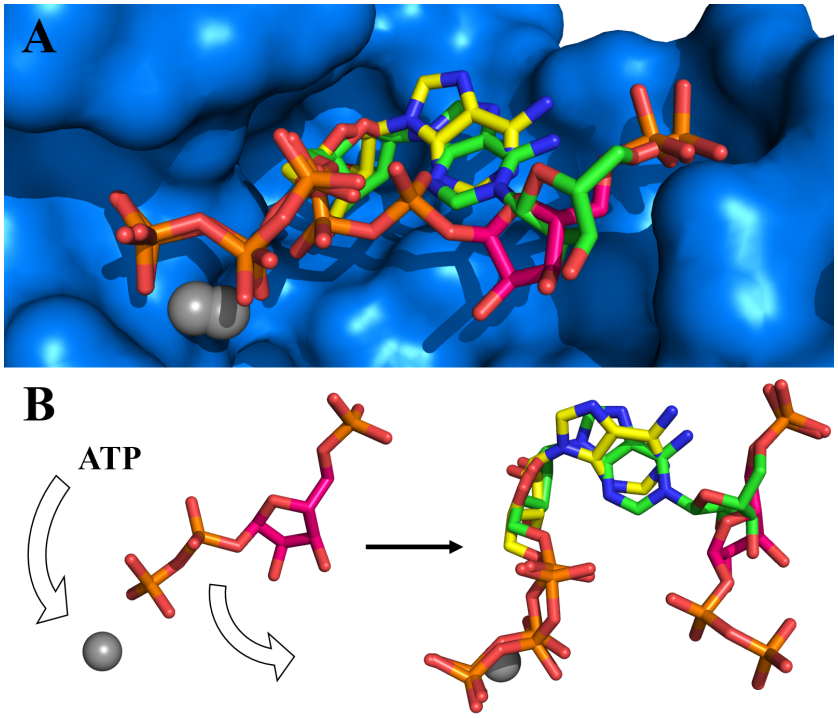


Figure 3.21: Occupation of the *Cje*ATP-PRT Core active site. A: ATP (yellow), PRPP (pink) and PR-ATP (green) represented as sticks as observed in the enzyme active site (blue), shown as surface. The three structures were superimposed on their α-carbon chains. R16 is removed for better visibility B: The two proposed binding states of PRPP in the active site. Individually PRPP is able to bridge over and coordinate to the bound metal. As ATP enters the active site, PRPP moves away and catalysis can occur, visualised by the product PR-ATP. Zinc is displayed as a sphere (grey). Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), phosphorous (orange).

The sensitivity of the *Cje*ATP-PRT Core to AMP is higher than for the wild type protein. ITC experiments also indicate a significantly higher affinity for ATP in the Core mutant. The substrate binding abilities are well conserved, evident by both ITC and crystallography, and can therefore be ruled out as the reason for the low catalytic turn over. It seems more likely that the high ligand affinities cause the enzyme to be kinetically trapped, making the dissociation of either ATP or PR-ATP the rate determining step. This assumption is strongly supported by both the determined product inhibition of PR-ATP and its high crystallographic resolution, as these indicate

a very slow release of the product from the active site. Overall it appears that the *Cje*ATP-PRT Core dimer is not well suited to conduct catalysis and merely provides a specific binding site.

The lack of the regulatory domain causes the catalytic domains to adopt a conformation that is seen for the dimeric units in all histidine bound long form ATP-PRT structures. SAXS analysis also confirmed the existence of a single species in solution, which fits well to the obtained *Cje*ATP-PRT Core crystal structure. The existence of the *Cje*ATP-PRT Core in this fixed, "naturally inhibited" conformation is likely to be the rationale for the low activity observed. The 3-fold increase in the K_M value for PRPP can also be accounted for by this fact, as the inhibited dimer interface causes changes in the putative pyrophosphate binding site.[36] Since *Cje*ATP-PRT Core predominantly adopts this inhibited conformation, the assumption can be made that it represents the lowest energy state or the most stable conformation for any ATP-PRT dimer. The decreased catalytic properties of the truncated long form mutants would therefore ultimately be caused by the disturbance of the hexameric state of the enzyme, which provides either conformational or dynamic control to the active sites by the way of ACT domain motion. It could be concluded that the ACT domains are required for *Cje*ATP-PRT to adopt the active and not the inactive state, and that they are as much required for effective catalysis as they are for feedback inhibition. This hypothesis opposes all previously drawn conclusions about the existence of an active dimeric long form ATP-PRT species,[26,35] as it is unlikely that the ACT domains on either end of the dimer cause significant changes in the dimer interface on their own.

While the vast majority of *Cje*ATP-PRT Core exists as a dimer, there is limited evidence for a higher order species with the size of a potential hexamer. Considering the distribution between the species by AUC, the higher order species can not account for the total *Cje*ATP-PRT Core activity measured, but its contribution can not be ruled out entirely. The observed peak may also correspond to a structurally undefined aggregate as a result of slow decay in solution.

In contrast to *Cje*ATP-PRT Core, the isolated *Cje*ATP-PRT ACT appeared to have conserved neither fold nor function of to the regulatory domain of the full length enzyme. Although it was expressed in a soluble form and likely adopted a trimeric quaternary structure, the *Cje*ATP-PRT failed to possess any measurable histidine binding abilities. Furthermore *Cje*ATP-PRT ACT was unable to interact with the *Cje*ATP-PRT Core in solution to provide for enhanced activity or histidine sensitivity. These results can be interpreted in different ways. The purified species could have been either misfolded or the assembly misarranged, which is strongly supported by the measured CD spectrum. Oligomeric state analysis on the other hand supports the presence of a defined soluble trimeric species. Considering its relatively high melting temperature, the isolated *Cje*ATP-PRT ACT species must have adopted a very stable and/or compact conformation. As crystallographic evidence is missing, the structural details of this conformation can only be speculated and it is questionable if it possesses any similarities to the ACT domain of *Cje*ATP-PRT. The majority of the ACT domain fold consists of β strands and a small shoulder around 220 nm was observed in the CD data. A possible conformation could therefore be composed of a triple central beta sheet, similar to the tight histidine bound state of the ACT domains in the *Cje*ATP-PRT wild type, with the surrounding helices being random coils. Whatever the actual adopted conformation is, the clear change in 3D structure and/or multimeric assembly is the likely reason for the loss of histidine binding abilities.

Apart from its role as an inhibitor binding module, which could not be proven by a direct investigation of the isolated species, the ACT domain trimer provides structural integrity to the hexamer. This is evident in the lower melting temperature (approximately 11 °C) of the Core mutant compared to the wild type *Cje*ATP-PRT.

By removing the ACT domain the *Cje*ATP-PRT was reduced to its catalytic apparatus only and it adopted a dimeric state, like isolated short form ATP-PRTs.[101] This unregulated species, *Cje*ATP-PRT Core, possesses strong binding abilities but lacks an efficient catalytic turn over. Following

the idea that both contemporary ATP-PRT subfamilies derive from the same ancestor,[37] the *Cje*ATP-PRT Core could be viewed to be a precursor-like intermediate species that could evolve in either direction, either by gene fusion with an ACT domain or by recruiting a secondary protein chain. The low catalytic capabilities of this intermediate clearly show that the modern long form ATP-PRT enzymes are not simply the sum of their structural and functional parts but require all domains and a complex quaternary structure for full functionality.

Chapter 4

The inhibitory signal transduction in *Campylobacter jejuni* adenosine triphosphate phosphoribosyl transferase

4.1 Introduction

Allosteric signal transduction, the way by which ligand binding to a remote site is communicated to the active site of an enzyme, occurs via a variety of different mechanisms. Depending on the enzyme, large or subtle changes in conformation or dynamics can occur, which rearrange domains, single structural elements, interfaces or simply side chains. Often these changes are communicated along a defined internal pathway by the formation or break down of key interactions.

The existence of a signalling pathway in *Cje*ATP-PRT that connects the remote allosteric histidine binding site to the active site, and vice versa, is conceivable due to the synergistic character of AMP and histidine inhibition. While the removal of the regulatory ACT domain crudely and undoubtedly

removed the regulation by the feedback inhibitor histidine, it also highlighted the need for the regulatory apparatus for the maintenance of the active state of the enzyme. It is clear that histidine binding causes a change in both the ACT domain interface, which is directly involved in the binding process, and the dimer interface, which influences the active sites. But other than the gross conformational changes, no internal pathway by which the signal is communicated has been identified.

An approach to segregate inhibitory response from enzymatic activity is the mutational exchange of residues that are positioned along the inhibitory pathway. If key interactions are broken or disturbed, the inhibitory transmission could be disrupted without affecting the enzyme function. In the following chapter approaches to delineate an exact pathway including the partial characterisation of one such mutant of *Cje*ATP-PRT are described.

4.2 Identification of potential residues of a inhibitory signal transduction pathway

The first step in the process of uncovering the signalling pathway was to identify key residues that may play a role in the signal transmission. A first idea about crucial interactions was obtained by inspecting the binding modes of the active site ligands presented in the crystal structures of the previous chapters. The search was focussed especially on changes that occur between the two observed conformations of *Cje*ATP-PRT.

Very pronounced changes were seen for the flexible α1-β2 loop, which forms the only inter-molecular connection in the active site of the AMP bound form (4YB6) via the side chains of the two histidine residues H33 and H35 (Figure 4.1). The loop occupies the same position in the histidine bound crystal structure (4YB5), while it is far away from any possible interaction in the ATP bound structure (4YB7) (section 2.9.6). It seems likely that this prearrangement of the loop is involved in the synergistic effect of histidine

on AMP inhibition and it was therefore hypothesised to be part of an inhibitory signalling pathway. Consequently, H33 and H35 were targeted by mutagenesis. Due to their close sequence proximity, it was decided to create a double mutant altering both residues simultaneously, H33 to alanine and H35 to serine. Serine was chosen to retain the hydrophilic character of the loop, while removing its binding ability.

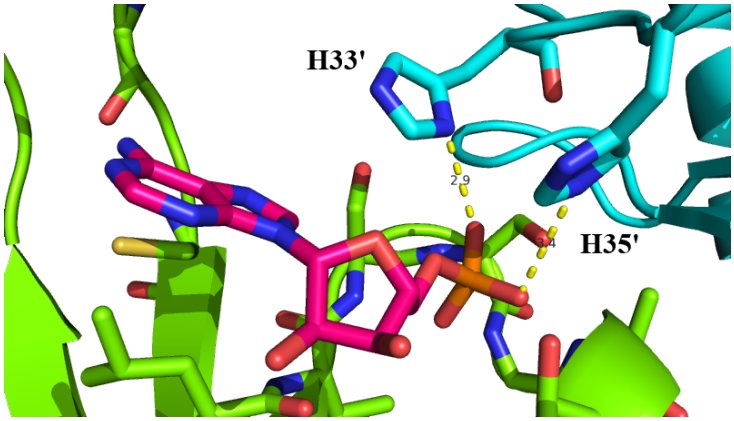


Figure 4.1: Double histidine motif of *Cje*ATP-PRT α1-β2 loop. Cartoon and stick representation of the *Cje*ATP-PRT active site. A single AMP (pink) is bound between chain A (green) and chain D (cyan). The two hydrogen bonding interactions with H33' and H35' are highlighted (yellow dashed lines). Residues in close distance to AMP are displayed as sticks. Atoms are coloured according to element: oxygen (red), nitrogen (blue), sulphur (yellow), phosphorous (orange).

Two more residues for mutation were chosen after careful inspection of the active and inhibited *Cje*ATP-PRT crystal structures. When superimposed on a single ACT domain, the relative position change of the surrounding structural elements became visible. One of the most prominent differences was found in the long α8 helix that connects domain I with the ACT domain of each chain. The helix exhibits a relative angle change of approximately 30° between the superimposed chains. Two arginine residues were identified that are seemingly involved in facilitating this movement, R216 and R224. Upon the movement of α8, R216 changes not only its position but also its connectivity (Figure 4.2). In the absence of histidine (4YB7) it interacts strongly only with W275. This interaction is also maintained in the presence of histidine (4YB5), but a new connection is formed to N272, bridging

the gap between the domains. This new connection was also one of the few interactions that does not directly involve the histidine binding site. R224 maintains its connectivity to D100 in both structures, but is turned sideways as the helix moves and the β5-β6 loop with D100 follows this motion. This movement was thought to have an influence on the directly adjacent active site. To abolish any possible side chain interactions, both arginine residues were mutated to alanine.

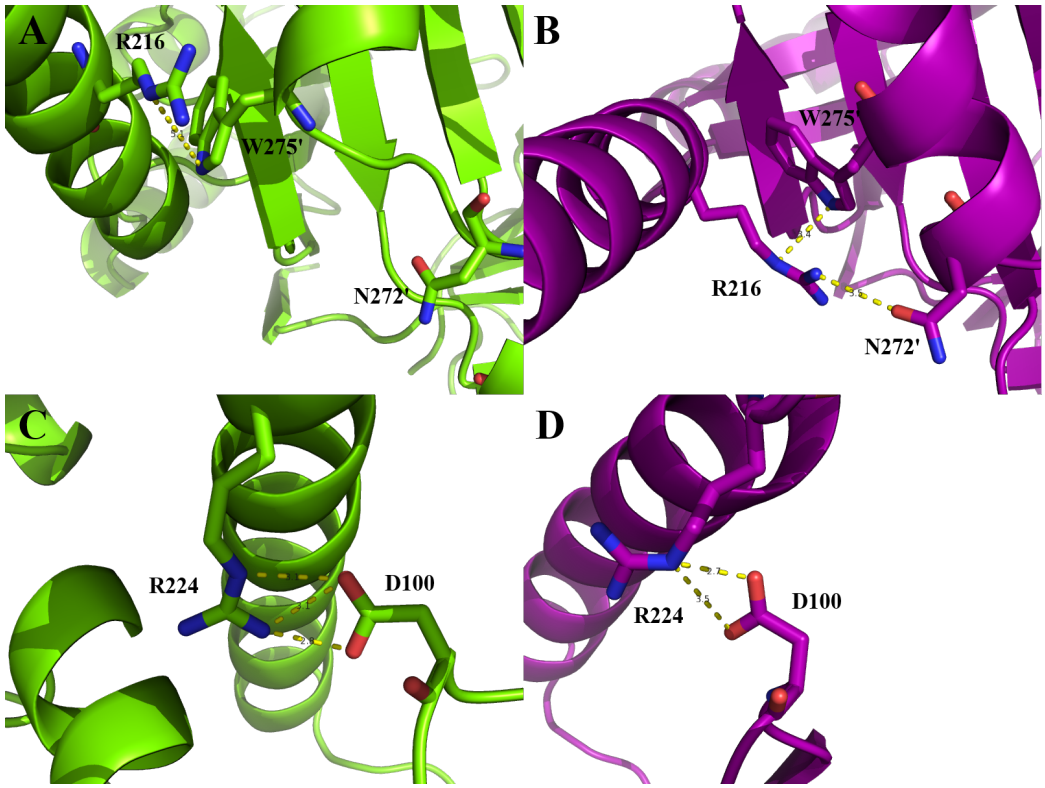


Figure 4.2: Relative positioning of the α8 residues R216 and R224. Environment of the arginine residues R216 (A and B) and R224 (C and D) in cartoon representation compared between the active (green) and inhibited (purple) *Cje*ATP-PRT crystal structures. Hydrogen bonding interactions are highlighted (yellow dashed lines). Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue).

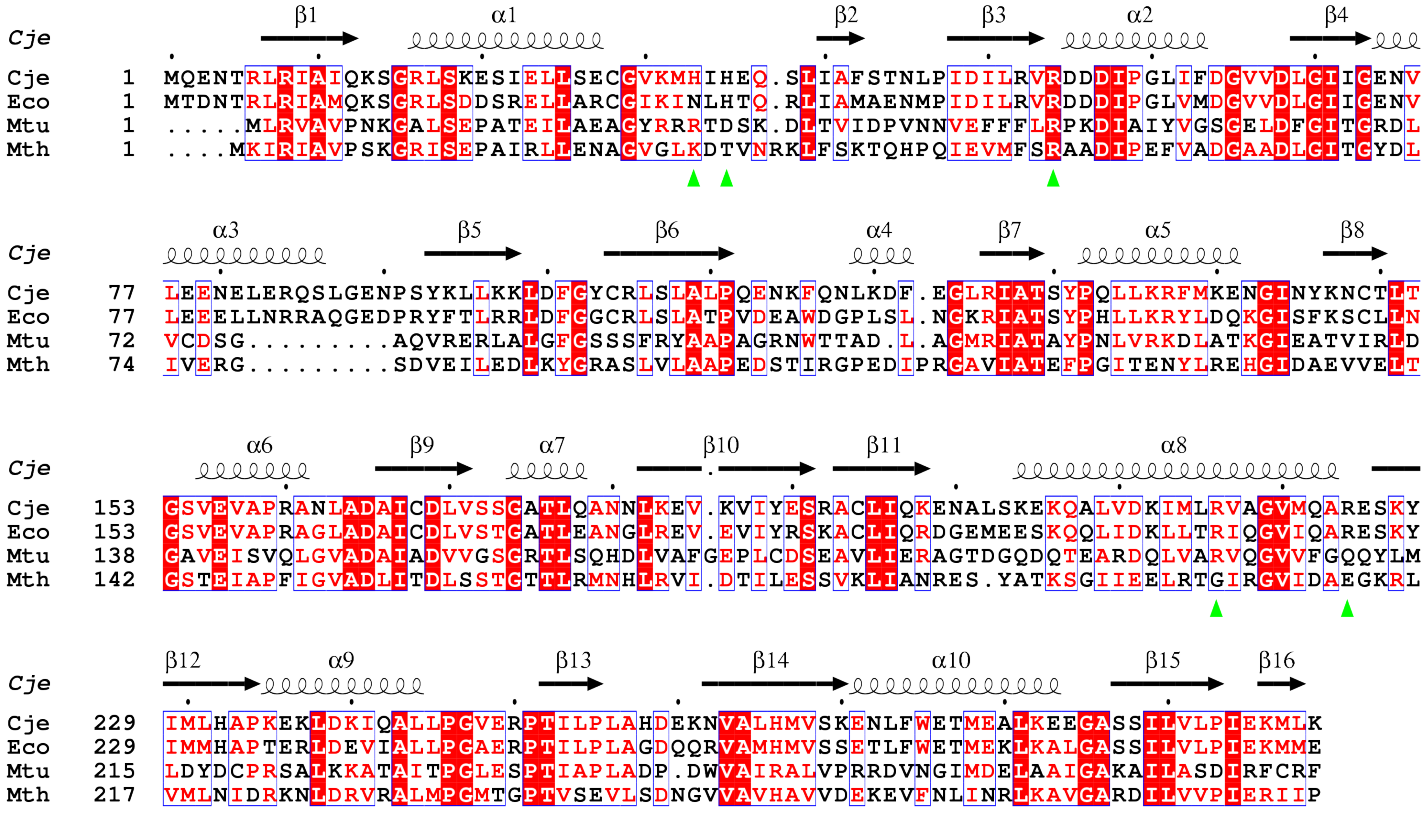
Histidine binding causes the dimer interface to undergoe significant changes, as described in chapter 2. The individual chains rotate relative to each other and form new contacts in close proximity to their active sites. One

of these contacts includes R54, a highly conserved residue that is responsible for the binding of the triphosphate group of ATP in the open *Cje*ATP-PRT conformation (4YB7). In the histidine bound crystal structure (4YB5) R54 is interacting closely with its symmetric counterpart R54' at the centre of the dimer interface. To disrupt a potential communication across the interface via this contact, R54 was chosen as a fourth target for mutagenesis. Like R216 and R224, R54 was mutated to alanine. Figure 4.3 shows the relative conservation of the chosen residues with respect to the structurally characterised long form ATP-PRT enzymes.

4.3 Cloning and purification

The site directed mutagenesis (procedure described in section 6.4.1) was successful for all three single site mutants, R54A, R216A and R224A, delivering sequence verified clones in pDEST^{TM}17 (Figure 4.4). While this approach also produced successful amplification of constructs for the double histidine mutant H33A/H35S it failed to generate the desired mutations. Multiple colonies from different mutagenesis approaches with different annealing temperatures were analysed by DNA sequencing without success. The mutations were either not present, or the mutagenesis primer was inserted into the gene multiple times (2x, 3x, and 4x). Amongst all sequenced clones, one contained both the parental plasmid with *Cje*ATP-PRT wild type sequence and a plasmid with the correct mutations, evident as a mixed read in the sequencing chromatogram. Because of these difficulties, especially the multiple primer

Figure 4.3 *(following page)*: Conservation of mutated *Cje*ATP-PRT residues. Multiple sequence alignment of all long form ATP-PRT sequences with available structural information. Secondary structure elements of *Cje*ATP-PRT are displayed above as arrows (β strands) and spirals (α helices). Residues are coloured according to conservation: absolutely conserved (white, red box), partially conserved (red) and not conserved (black). Green triangles indicate the position of the mutated residues. Organism three letter code: *M. tuberculosis* (Mtu), *M. thermautotrophicus* (Mth), *E. coli* (Eco) and *C. jejuni* (Cje). Alignment was created with ClustalOmega[84] and edited using ESPript.[88]


Cje
β1 α1 β2 β3 α2 β4
Cje 1 MQENTRLRIAIQKSGRLSKESIELLSECGVKMHIHEQ.SLIAFSTNLPIDILRVRDDDIPGLIFDGVVDLGIIGENV
Eco 1 MTDNTRLRIAMQKSGRLSDDSRELLARCGIKINLHTQ.RLIAMAENMPIDILRVRDDDIPGLVMDGVVDLGIIGENV
Mtu 1 .....MLRVAVPNKGALSEPATEILAEAGYRRRTDSK.DLTVIDPVNNVEFFFLRPKDIAIYVGSGELDFGITGRDL
Mth 1 ....MKIRIAVPSKGRISEPAIRLLENAGVGLKDTVNRKLFSKTQHPQIEVMFSRAADIPEFVADGAADLGITGYDL
α3 β5 β6 α4 β7 α5 β8
Cje 77 LEENELERQSLGENPSYKLLKKLDFGYCRLSLALPQENKFQNLKDF.EGLRIATSYPQLLKRFMKENGINYKNCTLT
Eco 77 LEEELLNRRAQGEDPRYFTLRRLDFGGCRLSLATPVDEAWDGPLSL.NGKRIATSYPHLLKRYLDQKGISFKSCLLN
Mtu 72 VCDSG.........AQVRERLALGFGSSSFRYAAPAGRNWTTAD.L.AGMRIATAYPNLVRKDLATKGIEATVIRLD
Mth 74 IVERG.........SDVEILEDLKYGRASLVLAAPEDSTIRGPEDIPRGAVIATEFPGITENYLREHGIDAEVVELT
α6 β9 α7 β10 β11 α8
Cje 153 GSVEVAPRANLADAICDLVSSGATLQANNLKEV.KVIYESRACLIQKENALSKEKQALVDKIMLRVAGVMQARESKY
Eco 153 GSVEVAPRAGLADAICDLVSTGATIEANGLREV.EVIYRSKACLIQRDGEMEESKQQLIDKLLTRIQGVIQARESKY
Mtu 138 GAVEISVQLGVADAIADVVGSGRTLSQHDLVAFGEPLCDSEAVLIERAGTDGQDQTEARDQLVARVQGVVFGQQYLM
Mth 142 GSTEIAPFIGVADLITDLSSTGTTLRMNHLRVI.DTILESSVKLIANRES.YATKSGIIEELRTGIRGVIDAEGKRL
β12 α9 β13 β14 α10 β15 β16
Cje 229 IMLHAPKEKLDKIQALLPGVERPTILPLAHDEKNVALHMVSKENLFWETMEALKEEGASSILVLPIEKMLK
Eco 229 IMMHAPTERLDEVIALLPGAERPTILPLAGDQQRVAMHMVSSETLFWETMEKLKALGASSILVLPIEKMME
Mtu 215 LDYDCPRSALKKATAITPGLESPTIAPLADP.DWVAIRALVPRRDVNGIMDELAAIGAKAILASDIRFCRF
Mth 217 VMLNIDRKNLDRVRALMPGMTGPTVSEVLSDNGVVAVHAVVDEKEVFNLINRLKAVGARDILVVPIERIIP

inserts, the initial primer design and preparation were replaced with a primer set containing a 7 base pair longer overlap on the forward direction 3'-end. Unfortunately this did not yield any suitable amplification under the conditions used either and this mutant was not pursued further due to time constraints.

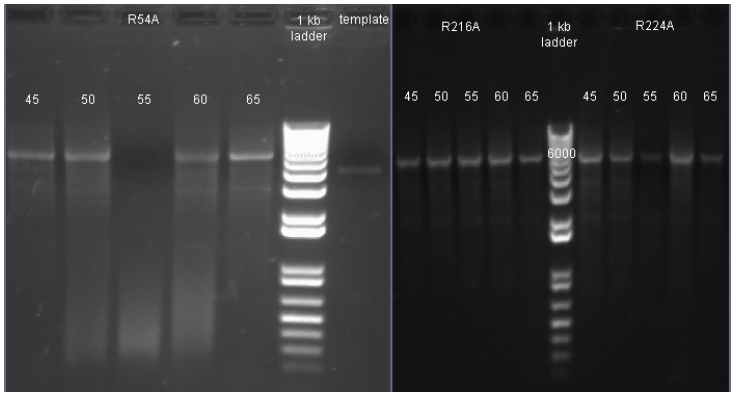


Figure 4.4: *Cje*ATP-PRT mutagenesis. Agarose gel electrophoresis of the site directed mutagenesis reactions generating the three single site mutants of *Cje*ATP-PRT. The reactions of R54A (left), R216A (middle) and R224A (right) were performed at an array of different annealing temperatures according to section 6.4.1 using *Pfu Ultra* High-Fidelity Polymerase. Electrophoresis was performed according to section 6.1.13

Test expression showed good solubility for all mutants in all strains tested. Consequently, all large scale expression and purification was performed with Chaperone3 strains, following the established protocol of the *Cje*ATP-PRT wild type. The purification results with respect to purity and yield were comparable with the wild type enzyme (Figure 4.5). As seen for the wild type, the highly pure mutant enzymes showed an additional band around 66 kDa corresponding to the dimeric size of *Cje*ATP-PRT, due to insufficient denaturation time.

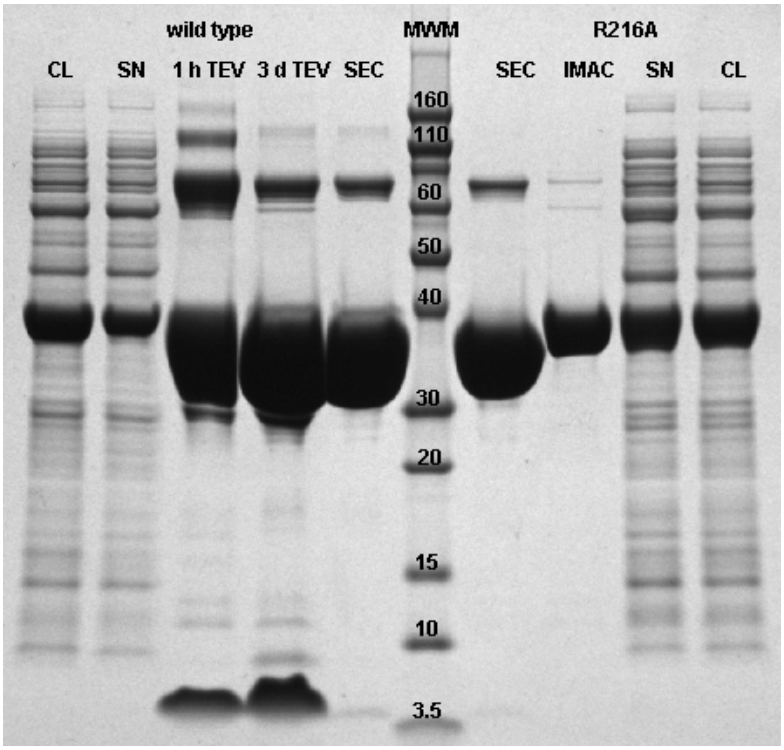


Figure 4.5: Purification of the *Cje*ATP-PRT R216A mutant. SDS-PAGE showing selected purification steps of both wild type and R216A enzymes in comparison. 15 μL samples were taken from the preparations at each step, denatured with 5 μL of sample loading dye for 2 minutes at 95 °C, loaded disregarding the differences in protein concentration. Gel was run according to section 6.1.31. Lane labels correspond to: MWM - molecular weight marker, CL - crude lysate, SN - clarified supernatant, IMAC - pooled eluate fractions of the IMAC column, TEV - sample during/after TEV protease incubation, SEC - pooled fractions after SEC.

4.4 Determination of molecular weight

The molecular masses of the purified *Cje*ATP-PRT R54A and R216A mutants were determined to be 33607.9 g mol^{-1} and 33609.4 g mol^{-1} respectively by electrospray ionisation mass spectrometry. These values are in agreement with the theoretical value of 33611.1 g mol^{-1} calculated for the 300 amino acid sequence of the wild type protein with a single arginine changed to alanine.

4.5 Mutant folding

To determine if the generated mutant enzymes adopted similar folds to the wild type enzyme, CD experiments were performed. All curves were similar in shape (Figure 4.6), showing the two main peaks for α helices (210 nm) and β strands (220 nm). None of the mutants showed signs of unfolding (200 nm) as the *Cje*ATP-PRT ACT protein did (section 3.9). These results are consistent with a similar secondary structure composition and native folding of the mutants analogous to the wild type protein.

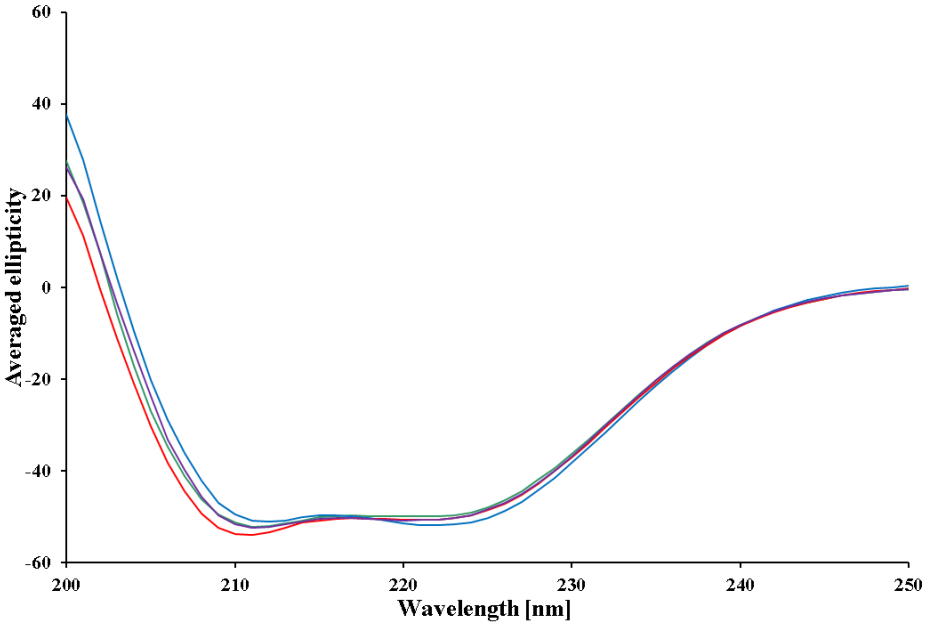


Figure 4.6: Folding of *Cje*ATP-PRT mutants. Five point averaged ellipticity was generated from the raw CD measurements according to section 6.1.36. The curves for *Cje*ATP-PRT wild type (green), *Cje*ATP-PRT Core (blue), *Cje*ATP-PRT R216A (red) and *Cje*ATP-PRT R54A (purple) are shown in comparison. Typically CD peaks are associated with: β strand (220 nm), α helix (210 nm) and random coil (200 nm).

4.6 Mutant activity

4.6.1 *Cje*ATP-PRT R54A

Neither crude nor highly pure *Cje*ATP-PRT R54A exhibited any measurable catalytic activity under standard assay conditions. Hence further kinetic characterisation of this mutant was not pursued. As the protein showed correct folding, the total loss of activity must be a direct consequence of the mutated residue. This is not completely unsurprising, as R54 plays a role in ATP binding (section 2.9.4).

4.6.2 *Cje*ATP-PRT R224A

The crude preparation (section 6.4.2) of the *Cje*ATP-PRT R224A mutant showed an activity of approximately 10 % compared to the wild type enzyme under standard assay conditions. It was also still responsive to histidine, being fully inhibited in the presence of 1 mM histidine (Figure 4.7). Due to this wild type-like behaviour it was excluded from any further characterisation.

4.6.3 *Cje*ATP-PRT R216A

Like the *Cje*ATP-PRT R224A mutant, the *Cje*ATP-PRT R216A mutant showed a largely reduced activity, compared to the wild type enzyme. No further activity loss was observed in the presence of 1 mM histidine though, a concentration at which the *Cje*ATP-PRT wild type is fully inhibited (Figure 4.7). This is a clear indication that the inhibitory response to histidine is compromised in this mutant. A comparison of the kinetic parameters determined for wild type and R216A mutant, as well as the Core mutant is summarised in Table 4.1. It is evident from these results that the catalytic properties of *Cje*ATP-PRT R216A are closer to the *Cje*ATP-PRT Core than to the wild type enzyme, with an increased K_M(PRPP) value and a two orders of magnitude lower k_{cat} value.

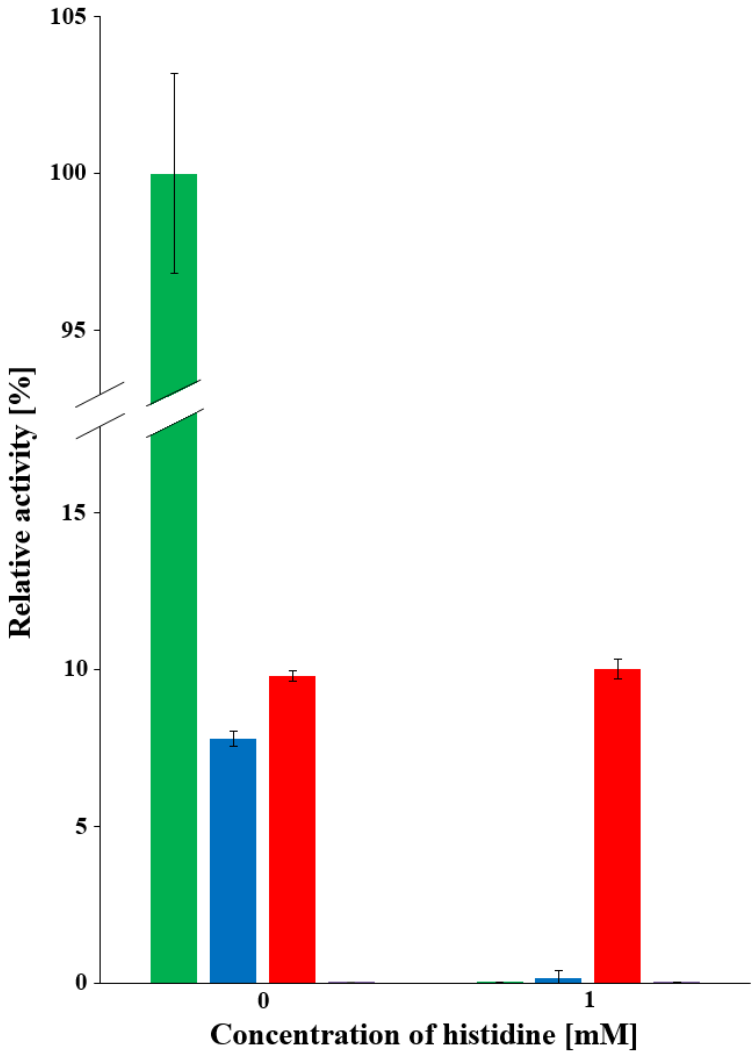


Figure 4.7: Comparison of the *Cje*ATP-PRT mutant activities. Specific activity was measured under standard assay conditions (section 6.1.33) using saturating substrate concentrations (2 mM ATP and 0.3 mM PRPP) in the presence of 0 and 1 mM histidine and normalised against the wild type value. The Y-axis break indicates a jump of 75 %. Colours correspond to: *Cje*ATP-PRT wild type (green), *Cje*ATP-PRT R224A (blue), *Cje*ATP-PRT R216A (red), *Cje*ATP-PRT R54A (purple).

To verify the initial results, the response of *Cje*ATP-PRT R216A to histidine was determined over three orders of magnitude (0.2–20 mM). The results show very low variance in the overall activity that does not follow a trend (Figure 4.8A), consistent with the absence of a histidine response.

Table 4.1: Kinetic constants of *Cje*ATP-PRT R216A. Measurements were performed according to section 6.4.3. Equivalent values of *Cje*ATP-PRT wild type and *Cje*ATP-PRT Core are given as a comparison.

*Cje*ATP-PRT	K_M(ATP) [μM]	K_M(PRPP) [μM]	k_{cat} [s^{-1}]
R216A	133 ± 19	36 ± 4	0.033 ± 0.001
Wild type	97 ± 10	15 ± 1	1.8 ± 0.1
Core	94 ± 11	41 ± 5	0.024 ± 0.003

The *Cje*ATP-PRT R216A mutant was furthermore found to be responsive to the active site inhibitor AMP. The IC_{50} of AMP was determined to 0.42 ± 0.05 mM and was found to be in good agreement with the IC_{50} value determined for *Cje*ATP-PRT Core, 0.29 ± 0.02 mM, while it differed significantly from the wild type value of 2.1 ± 0.3 mM (Figure 4.8B).

Like *Cje*ATP-PRT Core, the R216A mutant was also inhibited by low concentrations of the product PR-ATP. The IC_{50} of PR-ATP was determined to be 3.2 ± 0.3 μM, which is very similar to the value determined for *Cje*ATP-PRT Core, 6.7 ± 0.7 μM (Figure 4.8C).

These findings confirm that the R216A single site mutation only affects the inhibition by histidine. In contrast, the sensitivity and most likely the binding affinity towards active site ligands, in accordance with the results for *Cje*ATP-PRT Core, is increased. Overall the kinetic properties of *Cje*ATP-PRT Core and *Cje*ATP-PRT R216A are comparable, indicating that similar processes may have been compromised in both mutants.

4.7 Isothermal titration calorimetry

Histidine binding experiments were performed (section 6.4.4) on the *Cje*ATP-PRT R216A mutant to determine whether or not the observed lack of inhibition was the result of direct interference with the histidine binding site. The

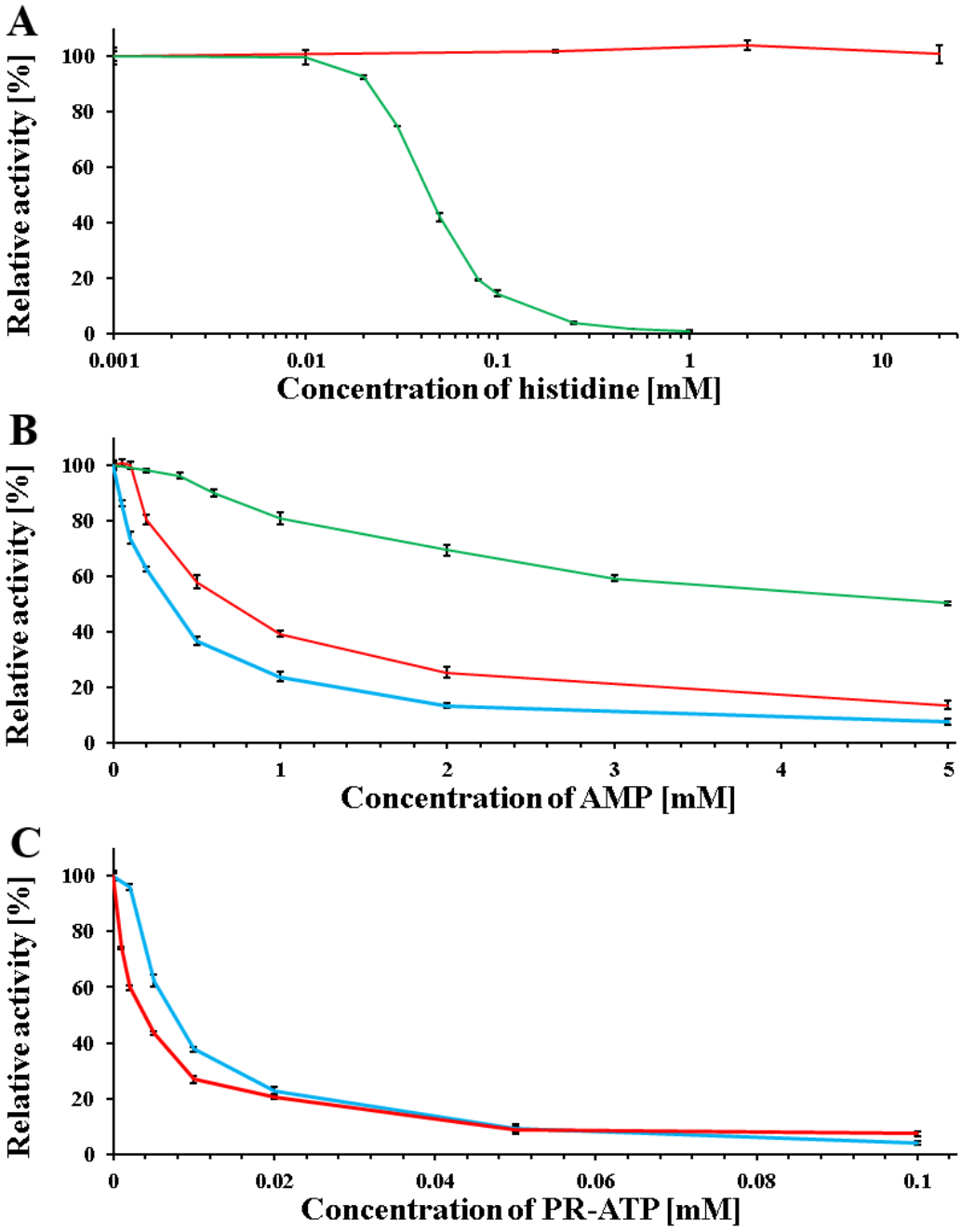


Figure 4.8: Inhibitory response of *Cje*ATP-PRT R216A. Dose response curves of *Cje*ATP-PRT R216A (red) to A: histidine, with the X-axis in logarithmical scale, B: AMP and C: PR-ATP. The response of the *Cje*ATP-PRT wild type (green) and *Cje*ATP-PRT Core (cyan) are given as a comparison.

resulting binding curve demonstrates that this is not the case. *Cje*ATP-PRT R216A is still able to bind histidine. (Figure 4.9).

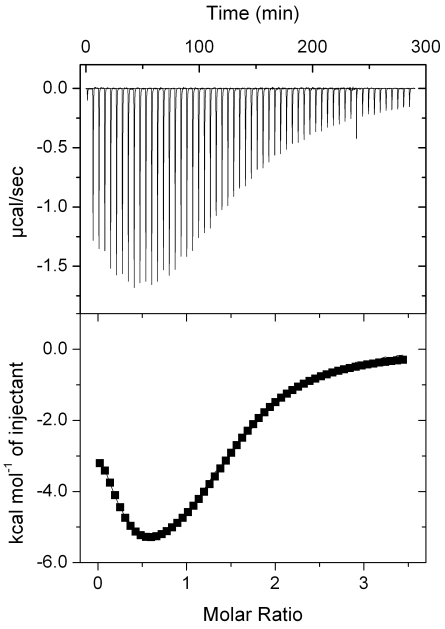


Figure 4.9: Histidine binding to *Cje*ATP-PRT R216A. Histidine binding curve was measured under the conditions described in section 6.1.39 and 6.4.4, subtracted by the corresponding heat of dilution data series and fitted using OriginLab® version 7.0.

The shape of the binding curve resembles the binding curve of histidine binding to the *Cje*ATP-PRT wild type in the presence of AMP (Figure 2.5D). The dissociation constant was determined by fitting the data to a sequential binding model with two binding events. The results presented in Table 4.2 are in good agreement with the K_d values measured for the wild type protein, clearly showing that the histidine binding site has not been affected. With both activity and allosteric binding confirmed, it is possible to conclude that the R216A mutation disrupts the transmission of the inhibitory response.

Table 4.2: Histidine dissociation constants. K_d values were determined by ITC according to section 6.2.4. A stoichiometry for the histidine titration was sometimes not determined (ND) due to the type of fit used.

*Cje*ATP-PRT	K_d [μM]	Stoichiometry
Wild type	19.3 ± 0.2, 14.5 ± 0.2	ND
AMP bound wild type	4.8 ± 0.7, 2.3 ± 0.1	0.85 ± 0.20
R216A	1.0 ± 0.3, 14.5 ± 0.7	1.28 ± 0.02

4.8 Crystallographic analysis

4.8.1 Crystallisation

Crystallisation was successfully achieved for the R54A and R216A mutants of *Cje*ATP-PRT in co-crystallisation conditions following the procedure described in section 6.4.5. Crystals formed in the presence of histidine, AMP and ATP individually for both mutants. All crystals had similar shapes but differed in size depending on the protein concentration and ligand used (Figure 4.10). Unfortunately all *Cje*ATP-PRT R54A crystals failed to diffract sufficiently for data collection up to this point.

Several datasets of *Cje*ATP-PRT R216A were collected featuring the different co-crystallisation conditions mentioned above. Overall the quality of the diffraction was poor (3.5–4.5 Å) Only two datasets of comparable resolution were chosen for further processing featuring two different co-crystals, the first one with histidine and the second one with ATP. The data was cut to 3.4 Å and integrated in space group $P12_11$. Employing a molecular replacement strategy, a single solution was found for both data sets with a search model generated from all six chains of the wild type crystal structure 4YB5. In both cases the initial solution contained the entire hexamer in the unit cell. Analogous to the search model, both structures were in the inhibited conformation. Later, during early refinement stages, the ATP co-crystal structure was not further pursued as the solution had obvious flaws regarding

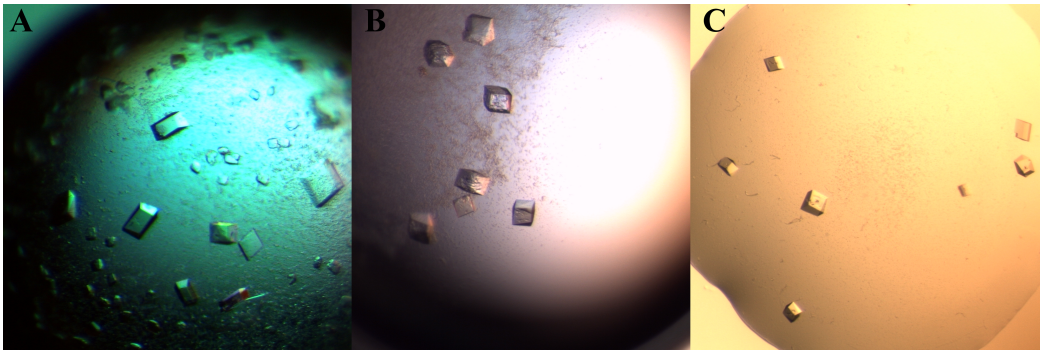


Figure 4.10: Crystals of *Cje*ATP-PRT R216A. Photographs taken of drops containing *Cje*ATP-PRT R216A crystals. Drop compositions were A: 0.1 M Tris/HCl, 0.2 M KSCN, 15–16 % w/v PEG 4000, 1 mM histidine, pH 7.5, B: 0.1 M Tris/HCl, 0.1 M $MgCl_2$, 13–15 % w/v PEG 4000, 5 mM AMP, pH 7.5, C: 0.1 M Tris/HCl, 0.2 M KSCN, 8 % PEG 1000/ 8 % PEG 8000, 10 mM ATP, pH 7.5.

density fit. The histidine dataset, on the other hand, was successfully refined (Table 4.3).

4.8.2 Comparison of overall fold

Compared to the *Cje*ATP-PRT wild type structure 4YB5, the *Cje*ATP-PRT R216A crystal structure was of more than 1 Å lower resolution. In all six refined chains, approximately 40 % of the amino acid side chains had to be removed, as well as entire residue ranges in areas of poor resolution. These areas mostly include the more flexible parts of domain II, α6 and the PRPP-binding loop (β9-α7), which underlines the importance of the mobility of these structural motifs for *Cje*ATP-PRT function.

No density was observed for the side chain of residue 216. Although the low resolution of the data set may have caused the side chain of arginine to be unresolved, the determined molecular weight for *Cje*ATP-PRT R216A is consistent with this observation. In some of the protein chains, the lack of interaction with the arginine also led to a visible reorientation of the tryptophan W275 side chain in the neighbouring subunit, which is another indication for the presence of the single site mutation.

The six chains in the unit cell arrange into the closed or inhibited hexamer conformation, as seen for the wild type structures 4YB5 and 4YB6, with the typical tight interactions between the ACT domain trimers. The dimer interface hereby buries a surface of 529.5 Å^2, approximately 4.7 % of the total surface area of one chain, and the trimer interface 855.8 Å^2, equivalent to 6.2 % total chain surface area. The overall lower surface area, compared to the wild type structures (section 2.9.7), is a direct consequence of the missing side chain interactions due to the low resolution (3.4 Å). The closed conformation of this structure becomes most clear by inspecting the superimpositions with the open and closed wild type crystal structures, with overall RMSD values of 1.16 Å to 4YB5 and 1.30 Å to 4YB6, but 40.6 Å to 4YB7 (Figure 4.11).

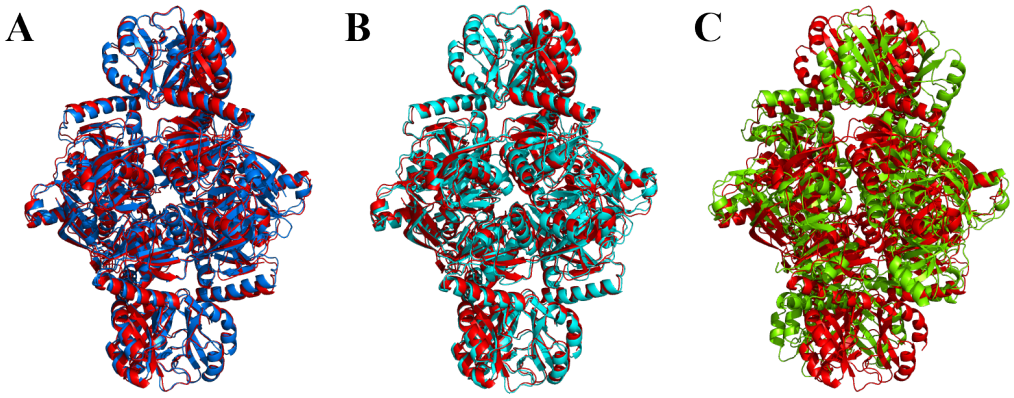


Figure 4.11: Conformation of histidine bound *Cje*ATP-PRT R216A. Superimposition of the *Cje*ATP-PRT R216A crystal structure (red) with the three wild type structures 4YB5 (blue, A), 4YB6 (cyan, B) and 4YB7 (green, C) in the cartoon representation.

4.8.3 Histidine binding

Ligand density was observed in all six chains of *Cje*ATP-PRT R216A at the histidine binding site. Although the resolution was not sufficient to clearly identify the bound molecule, it was assumed, based on the apparent site specificity, that it is histidine (Figure 4.12C). As the ligand density was a lot less defined, the binding modes of the modelled histidine molecules differ, but

they appear to consistently form the same contacts as observed in the wild type structures (Figure 4.12). The interactions with H226, Y228 and T252 were not observed in the mutant crystal structure, which is most likely due to the conserved water molecule not being resolved. The presence of histidine in the overall closed enzyme conformation is somewhat unsurprising, but supports the previous findings on the conserved histidine binding function of the *Cje*ATP-PRT R216A ACT domains.

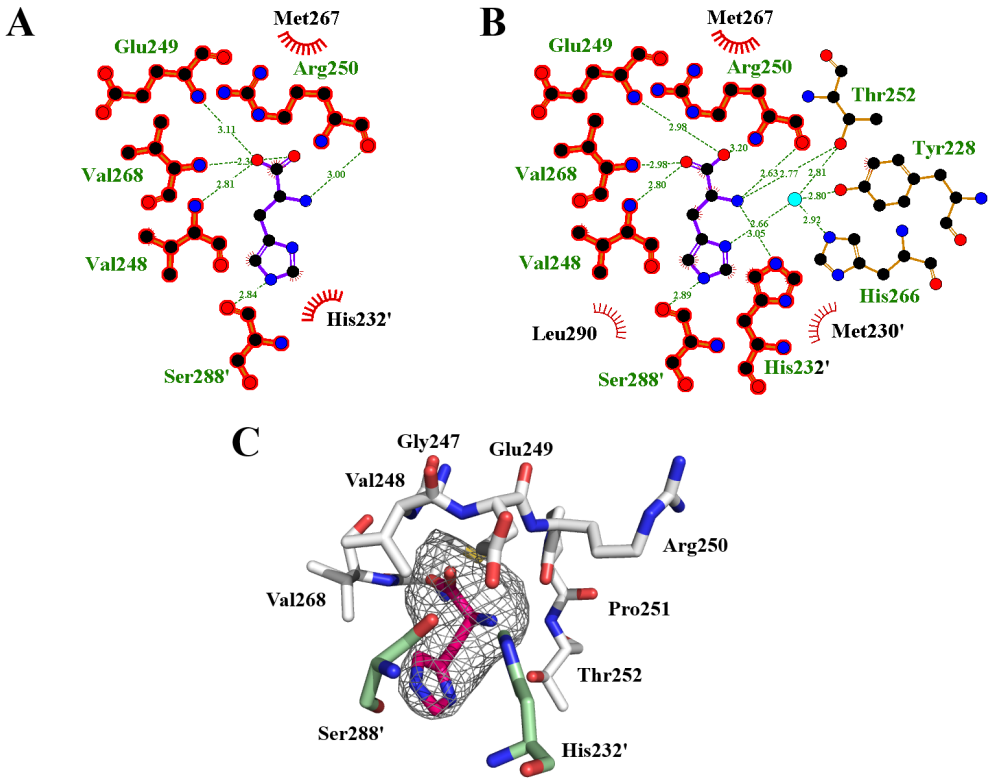


Figure 4.12: Histidine binding contacts. Observed binding mode of histidine. A and B: Planar plots of all observed histidine interactions in the *Cje*ATP-PRT R216A (A) and *Cje*ATP-PRT wild type (4YB5) (B) crystal structures created with LigPlot^{+}.[94] Hydrogen bonds are depicted as dashed lines labelled with the real space distances between atoms. Hydrophobic interactions are displayed as short radial red lines. Water molecules are displayed as cyan spots. Conserved interactions are highlighted in red. C: Stick representation of the histidine (pink) between chain E (white) and chain D (pale green) in *Cje*ATP-PRT R216A surrounded by the observed electron density (Fo-Fc map - grey mesh). Hetero-atoms are coloured according to element: oxygen (red), nitrogen (blue), sulphur (yellow).

Table 4.3: Preliminary crystal parameters, data collection, and refinement statistics for the histidine bound *Cje*ATP-PRT R216A structure.

	*Cje*ATP-PRT R216A
Data collection	
Crystal system	monoclinic
Space group	$P12_11$
Unit cell parameters	
a, b, c [Å]	92.09, 122.52, 99.31
α, β, γ [°]	90.00, 107.84, 90.00
Resolution range [Å]	48.23–3.40 (3.60–3.40)
Measurements	170172
Unique reflections	28893
Completeness	99.4 (98.0)
Redundancy	5.9 (5.9)
$I/\sigma(I)$	9.9 (1.9)
R_{merge}	0.122 (0.918)
$CC_{1/2}$	0.998 (0.699)
Matthews coefficient	2.64
Refinement	
Resolution [Å]	50.00–3.40
R_{cryst}	0.262
R_{free}	0.325
Chain length	300
Observed residues	289, 285, 284, 281, 274, 288
Water molecules	0
Others	6 histidine
Mean B [Å^2]	
Protein	94.00
Ligand	83.32
RMSD	
Bond lengths [Å]	0.011
Bond angles [°]	1.51
Chiral volumes [Å^3]	0.074
Ramachandran	
Preferred %	96.55
Allowed %	3.45
Outliers %	0.00

4.9 Discussion

The results presented in this chapter demonstrated the uncoupling of the allosteric inhibition and the enzymatic activity of *Cje*ATP-PRT by mutating a single residue, arginine R216. This mutation is centred in a pivotal point of the motion of the regulatory ACT domain. As already observed for the *Cje*ATP-PRT Core mutant, the interference with the natural movement of the regulatory apparatus caused a heavy catalytic penalty (approximately 90 % activity loss). In turn the sensitivity to the feedback inhibitor histidine was completely removed, although histidine was evidently shown to still bind to the enzyme.

The kinetic properties exhibited by *Cje*ATP-PRT R216A were very similar to the properties of *Cje*ATP-PRT Core. The measured K_M and k_{cat} values as well as the IC_{50} values of AMP and PR-ATP and the lack of any detectable histidine inhibition are all consistent with this conclusion. In the previous chapter it was argued that the Core mutant exists permanently in the inhibited conformation, as determined by its crystal structures. Following from this idea, it seems likely that the majority of the R216A mutant hexamers also adopt the inhibited conformation, which would explain the similarity in the kinetic behaviour of the two mutants. In contrast to the *Cje*ATP-PRT Core, crystallographic evidence for this assumption is weak for *Cje*ATP-PRT R216A. The two processed low resolution data sets were in fact both solved with the closed hexamer conformation as a search model. Only the histidine bound structure was refined though and shows the closed conformation, as expected. At 3.4 Å resolution most molecular details are missing, but the overall conformations of the open and closed wild type hexamers are quite different, which makes this assumption possible. More conclusive evidence in form of an ATP bound structure would be beneficial to clarify this hypothesis.

The reason for the poor diffraction quality of the mutant crystals is not entirely clear. All tested crystals showed a certain degree of fragmentation during cryoprotection and it is therefore likely that the cryoprotection

strategy used was not appropriate. Working with the wild type *Cje*ATP-PRT crystals always posed difficulties because of the relatively high solvent content. Together with the age of the protein solution used for crystallisation, this could also be a reason for the poor diffraction. Focussing on the optimisation of these final steps should resolve the issue and result in higher resolution data sets for the crystallised mutant proteins.

Clear evidence for the preserved histidine binding abilities of the *Cje*-ATP-PRT R216A ACT domain was provided by ITC and to a lesser extent by crystallography. The solved crystal structure supports the presence of a small molecule in the histidine binding site of all six chains. Given the presence of 1 mM histidine in the crystallisation condition and the low determined K_d values for histidine (1 and 15 μM), this ligand is most likely histidine. Additional candidates include Tris, SCN^-, 1,2-ethylendiol, and possibly phosphate or imidazole from previous purification steps, but none of these molecules has ever been observed to bind the site. The ITC binding curve of the R216A mutant has the cooperative shape seen for histidine binding to the wild type enzyme, but lacks the strong initial exothermic signal. Assuming that this early part of the curve is the result of the expected conformational change of the *Cje*ATP-PRT wild type hexamer, the loss of it may be interpreted as the loss of the conformational change. This is an indication that the binding of histidine occurs, but has no effect on the enzyme conformation.

Examination of the *Cje*ATP-PRT wild type open and closed conformations revealed a distinct difference in the connectivity of two neighbouring ACT domains involving a newly formed hydrogen bonding interaction between R216 and N276' (and possibly E271') in the inhibitor bound conformation. With the binding of histidine established in the R216A mutant and the missing inhibitory response, it is clear that the communication of the inhibitory signal flows along this intermolecular inter-domain contact, underlining the importance of the hexameric quaternary structure for full enzyme functionality. In the absence of this contact the enzyme behaves as if the ACT domain has no regulatory function, similar to its removal.

How the inhibitory signal is further communicated can only be speculated at this point. If the inhibitory effect is indeed largely based on conformational changes around the active site, it is most likely that the motion of the ACT domain leads to the consequent change of the dimer interface arrangement, as seen in the wild type crystal structures. It is also possible that the inhibitory response changes the dynamics of the active site, impeding the necessary motions for effective catalysis. Indications that this might be the case came from the thermal stability analysis of the wild type enzyme. The inhibitor bound *Cje*ATP-PRT resisted higher temperatures than unliganded samples, which is commonly correlated with a higher rigidity and lower dynamic freedom. Equivalent studies on the R216A mutant should be undertaken to support this theory. Whatever the exact mechanism of active site inhibition may be, the results presented indicate that the effect of the introduced mutation is based around the arrest of hexamer motion, causing the *Cje*ATP-PRT R216A mutant to largely exists in a low energy, "closed" conformation. Though *Cje*ATP-PRT R216A is capable of efficient histidine binding, it has no further effect on the activity. This supports the conclusion, that the ACT domain and consequently the hexameric arrangement of the enzyme is crucial not only for the induction of the inhibitory response, but also for the maintenance of an active enzyme state.

In contrast to the *Cje*ATP-PRT R216A mutant, the R54A and R224A mutations did not provide any further insight into the inhibitory signal transduction. The removal of the highly conserved R54 side chain seems to be detrimental for the *Cje*ATP-PRT activity, most likely due to its role in the binding of the ATP triphosphate moiety. This assumption is supported by the wild type-like folding of the R54A mutant, but could not be explored further due to the lack of crystallographic evidence. *Cje*ATP-PRT R224A showed comparable activity levels to *Cje*ATP-PRT R216A, but no directly observed influence on the histidine inhibition. The differences in histidine sensitivity have not been examined in detail though, leaving the possibility that this mutant might be receptive to much lower histidine concentrations. The salt bridge formed between R224 and D100 is observed in both, the

active and the inhibited complex, although their environment moves significantly. R224 could therefore simply posses a structural function, anchoring the ACT domain hinge, on which it is positioned, to the rest of domain I, bestowing stability on the enzyme. A loss of this stabilising effect may have been the reason for the low catalytic activity.

Overall, this chapter has provided evidence that the regulatory signal transduction in *Cje*ATP-PRT can be perturbed by targeting single amino acids. Furthermore it appears that the inflicted penalties on the catalytic properties of the enzyme are closely linked to the regulatory domain and its movement, as hypothesised for the *Cje*ATP-PRT Core truncation mutant.

Chapter 5

Summary

5.1 Active hexameric state of *Cje*ATP-PRT

The results presented in this thesis are consistent with the existence of an active, homo-hexameric long form ATP-PRT species originating from *C. jejuni*. Previous studies focussing on the crystal structures of *Eco*ATP-PRT and *Mtu*ATP-PRT had hypothesised that these enzymes exist in an equilibrium between active dimers and inhibited hexamers.[26,35] This hypothesis shows strong parallels to the morpheein model of allostery, where a monomeric state equilibrium determines the association into functionally different oligomers. But this is not the case for *Cje*ATP-PRT, which has been found to exist entirely as a hexamer by all used methods even at concentrations as low as 0.025 mg mL $^{-1}$ (742 nM monomeric concentration). Therefore the observed activity must be exhibited by the hexameric species.

As reported for *Eco*ATP-PRT[26] and *Mtu*ATP-PRT,[35] the crystal structures of *Cje*ATP-PRT show the enzyme in two distinct hexameric conformations. It is important to compare the entire hexamer, as the details of the conformational change lie mostly in the relative positioning of the six chains to each other and only limited information can be gathered from a direct superimposition of single chains. The active state, represented by the ATP

bound structure 4YB7 differs from the inhibited state, represented by the structures 4YB5 and 4YB6, by an approximately 30 ° twist of the regulatory domain in each chain and by a flex of each individual dimer interface and the resulting rearrangement of secondary structure elements. All changes taken together, the active state appears more loose, shorter and wider than the inhibited form, which is stretched, compact and tight in comparison.

The trimer interface that is formed by the interlocked β-sheets of three ACT domains was found to be substantial in all observed crystal structures of *Cje*ATP-PRT, equally important or even stronger than the functional dimer interface. This is in accordance with the high melting temperature found for the *Cje*ATP-PRT ACT protein on its own. Consequently, the removal of the ACT domain caused the disruption of the hexameric enzyme complex. The product of this truncation, *Cje*ATP-PRT Core, exists as a functional homo-dimer that is not feedback inhibited by histidine, but has also greatly impaired catalytic properties. It is arguable if this impairment is caused by the direct interference with the protein chain or the change in oligomeric state, but all solved crystal structures of this mutant show the dimer in a conformation very close to the histidine bound conformation of the wild type protein. Given the consistently low binding constants and clear ligand density in the *Cje*ATP-PRT Core, a lowered substrate affinity can be excluded as a potential reason for the activity loss. It may be possible that the inhibited state of the enzyme binds the substrate and/or product molecules too tightly, making the release of the products the rate-determining step. This theory is supported by the introduced product inhibition to the *Cje*ATP-PRT Core enzyme and its PR-ATP bound crystal structure.

The existence of the ACT domain and its natural conformational motions in *Cje*ATP-PRT appear to be crucial not only for the conformational changes imposed on the enzyme during histidine binding, but also for the preservation of the enzyme activity in the absence of histidine, which is indicated by kinetic results from the *Cje*ATP-PRT Core and *Cje*ATP-PRT R216A mutants. This suggests a much more important role of the ACT domain for the enzyme than simply providing inhibitor binding properties,

making the hexameric quaternary structure of the enzyme essential for regulation and effective catalysis.

5.2 Evolution of the ATP-PRT enzyme

The creation of the active truncated dimeric *Cje*ATP-PRT brought with it evolutionary implications. The truncation transformed the long form chain into a ATP-PRT "short form-like" protein. While similar in chain length and composition, as well as secondary and tertiary structure, it does not contain the specific sites required for the interaction with HisZ, the essential second component of functional short form holo-enzymes. The *Cje*ATP-PRT Core forms a dimer that resembles the two catalytic dimer units present in short form ATP-PRT enzymes. These central units are the common feature of both ATP-PRT forms and it is likely that the large contemporary homo-hexameric and hetero-octameric complexes have evolved from a common ancestor.[37]

The recruitment of a member of the P_{II} signalling protein family and later gene fusion seems to be the most likely event that created the long form ATP-PRT enzymes.[48] A very similar, but not covalent, interaction with a P_{II} protein was identified to exist for *N*-acetyl-L-glutamate kinase, which represents a similar evolutionary development to the formation of the long form ATP-PRT enzymes.[49] The short form ATP-PRT complexes, on the other hand, are believed to have developed from the association with a histidyl-tRNA synthetase gene product that originated from a gene duplication event and later evolved into the HisZ protein.[45] Extensive phylogenetic studies showed that HisZ only exists in genomes that encode for a short form ATP-PRT gene, underlining the tight evolutionary relationship of those two genes.[51] The common ancestor of both ATP-PRT forms was most likely also a catalytic dimer, as this structure has been preserved throughout evolution. The artificial creation of *Cje*ATP-PRT Core could therefore be regarded as a successful evolutionary regress into a ancestor-like state.

The very slow catalytic turn over of the *Cje*ATP-PRT Core implies that the catalytic properties of the ancestral ATP-PRT enzymes could have been comparably poor as well. The development of large complexes consequently suited two purposes, an improvement on the catalysed reaction kinetics and the installation of the feedback inhibition by the end product of the (evolving) pathway. Essential for this development was the introduction of a mechanisms by which the active sites of the dimer could be altered between an active and inhibited state[27,36] and the coupling of this change to effector binding. The recruitment of the regulatory domain by the later long form added a structural element for histidine binding, the ACT trimer. This led to the formation of the hollow hexameric ellipsoid structure, which was then able to modulate the contained dimers by twisting and stretching (Figure 5.1A). A similar process can be imagined for the octameric short form complex. A change in the HisZ conformation at the central hinge point, widening the angle between the two HisZ dimers like scissors, would cause a rotational stretch of the peripherally bound ATP-PRT dimers (Figure 5.1B). Following this hypothesis both parallel evolutionary paths led to similar result using different molecular mechanisms.

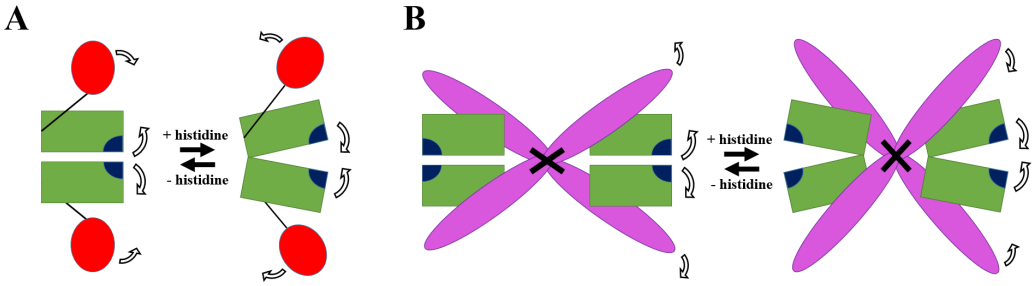


Figure 5.1: Mechanisms of dimer rearrangement. Schematic depiction of the observed and hypothetical movements of the catalytic ATP-PRT dimer units (green) facilitated by their corresponding regulatory apparatus (red and pink). A: The catalytic sites (blue) are arranged almost parallel to each other in the active conformation of the long form ATP-PRT. Histidine binding causes a twist of the ACT domains (red) to form the tight trimer interface (not shown), which pulls the active sites in the catalytic core apart. B: A widening of the angle between the HisZ dimers (pink) at their hinge point (black cross) might cause similar changes to occur in the short form ATP-PRT dimers.

5.3 The inhibitory signal transduction

Mutational analysis of *Cje*ATP-PRT identified one particularly interesting mutation site, arginine R216. The mutation to alanine resulted in a mutant enzyme that lacks the allosteric regulation by histidine. The mutated residue is positioned at the interface between helix $\alpha 8$, of its own chain, and helix $\alpha 10$ of the neighbouring chain. This interface undergoes significant changes upon binding of histidine, moving closer together and forming new contacts including hydrogen bonding interactions with R216. The lack of these contacts in the R216A mutant has been shown to not adversely affect histidine binding. R216 must therefore play a crucial role in the molecular communication of the inhibitory signal.

The kinetic properties of the *Cje*ATP-PRT R216A and *Cje*ATP-PRT Core mutants are in very close agreement. The very similar catalytic turn over rate and the strongly increased sensitivity to AMP and the reaction product PR-ATP are indicative of this fact. The binding of histidine to the R216A mutant was confirmed by ITC, but lacks the initial strong exothermic signal observed for the wild type, which may be associated with a rearrangement of the enzyme complex altering the properties of the histidine binding sites. This conformational or energetic change is absent in *Cje*ATP-PRT R216A. Preliminary crystallisation results also indicate that the R216A mutant structure might adopt the histidine inhibited conformation in the presence of ATP, unlike the wild type enzyme. Taken together these findings imply the existence of a permanently or predominantly inhibited conformation of the catalytic core of *Cje*ATP-PRT R216A, as seen for the *Cje*ATP-PRT Core dimer. It is therefore hypothesised that the mutation of R216 has detached the ACT domain motion from the catalytic dimer motion, which effectively broke the transmission of the regulatory signal along this pathway resulting in a fixed "false" signal. The catalytic core is receiving a constant histidine binding signal, resulting in permanently reduced catalytic efficiency, while the "disconnected" ACT domain still allows for effect-less histidine binding.

With only limited crystallographic evidence from the generated mutants, little can be inferred about the exact nature of the ATP-PRT inhibition. The AMP bound structure of *Eco*ATP-PRT (1H3D) clearly adopts the open hexamer conformation, leaving doubts about whether the inhibitory effect is solely conformational. The inhibitors AMP and histidine both increase the T_m of *Cje*ATP-PRT, while the substrate ATP does not. This leads to the conclusion that the inhibitory signal rigidifies the complex, which supports a change in dynamics. More evidence was brought about by the direct comparison of the ATP bound structures of *Cje*ATP-PRT wild type and Core. The conformations of the single chains are very similar and ATP adopts the same binding mode, but the activity of both species is undoubtedly different. This is indicative of the alteration of essential enzyme dynamics by the removal of the ACT domain and consequently also by the break of the R216 contact point.

However, the crystallographic observations on the wild type enzyme support a conformational driven regulatory signal transmission, which involves large domain movements that are communicated via hinge points. These movements in turn force the dimer interfaces into different arrangements altering the catalytic properties of the ATP-PRT active sites. R216 sits on one of the crucial connection points between domains and a hypothetical communication pathway focussing on similar connections is shown in Figure 5.2). The pathway naturally includes the residues of the histidine binding site in the ACT domain and travels along inter-subunit interactions.

The α1-β2 loop is thought to be another important connection point. In the histidine bound, closed conformation the loop protrudes into the active site of a neighbouring chain. Therefore the loop may serve as a "sensor" for the binding of the active site inhibitor AMP. This interaction could be the reason for the increased AMP binding affinity of the wild type enzyme in the presence of histidine and thereby explain the synergistic inhibitory effect. Regarding the effects of AMP binding on histidine binding, the presence of AMP enables loop interactions, which might lower the energy required for the conformational change of the hexamer. The C-terminal side of the α1-

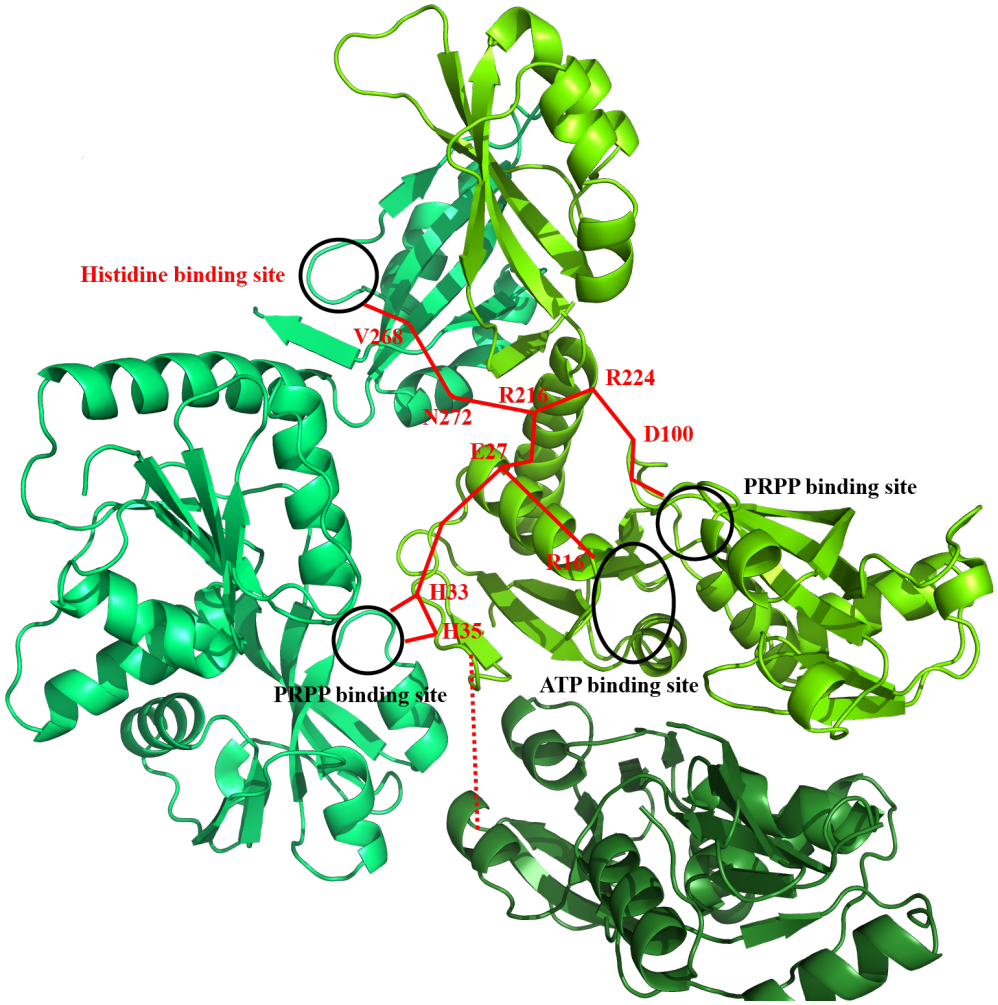


Figure 5.2: Hypothetical inhibitory signal communication pathway. Cartoon representation of three adjacent chains in the *Cje*ATP-PRT crystal structure 4YB6 with the third chain (dark green) only partially shown. A schematic pathway (solid red lines) between the allosteric and active sites (black circles) is superimposed onto the three dimensional image, highlighting potential key contacts in the transmission of the regulatory signal. The dashed red line depicts the lost interactions between the α1-β2 loop and helix α7 at the dimer interface.

β2 loop in turn is part of the open dimer interface interacting closely with residues of helix α7. A connection that is completely removed in the closed conformation. This might also direct the inhibitory signal flow from the dimer interface to the active site.

This is not the first report of uncoupled activity and feedback regulation in a long form ATP-PRT. The complete truncation of the regulatory domain of *Cgl*ATP-PRT generated a mutant that is not responsive to histidine and also has a reduced specific activity. These results are largely consistent with the findings on the *Cje*ATP-PRT enzyme presented in this thesis and validate that similar mechanisms for the inhibitory signal transmission exist in all long form ATP-PRT enzymes. Furthermore the debilitating effect of the ACT domain removal was able to be partially compensated by a single site mutant, S143F. This mutation did not only increase the specific activity of the *Cgl*ATP-PRT truncation mutant, but it also removed the feedback inhibition of histidine in the full length *Cgl*ATP-PRT enzyme.[98] The site of this randomly generated mutation, and the equivalent *Cje*ATP-PRT residue A158, is positioned in the α6 helix at the dimer interface. From a crystallographic view point this mutation site has no obvious involvement in any regulatory pathway other than its position close to the dimer interface. The change to a large hydrophobic side chain most likely disrupts the positioning of the two chains relative to each other. This leads to the assumption that this conformational change positively alters the catalytic abilities of the enzyme by "fixing" the dimer interface in the active conformation, directly opposite to the effect caused by the *Cje*ATP-PRT R216A mutation.

Overall the presented results strongly support a simple physical mechanism for the allosteric regulation of long form ATP-PRT enzymes by histidine. Two different hexamer conformations exist that differ greatly in their exhibited catalytic properties. In the absence of histidine, the equilibrium between the two states favours the active conformation, whereas histidine binding stabilises the inactive conformation and drives the equilibrium in the other direction. Due to the architecture of the enzyme and the nature of the ACT domain interfaces, the conformational change always affects the entire hexamer, with all chains adopting the same conformation. These described characteristics are consistent with the concerted MWC model[67] of allostery.

5.4 Future work

5.4.1 What is the nature of the PRPP binding?

Contrary to previous reports, PRPP was shown to be able to bind *Cje*ATP-PRT in the absence of ATP. Unfortunately, the analysis of the active site binding abilities is limited by the equilibrium character of the ATP-PRT reaction. The use of substrate mimics such as the carboxylic PRPP analogue 1-α-pyrophosphoryl-2-α,3-α-dihydroxy-4-β-cyclopentane-methanol-5-phosphate (PCP) could be very beneficial to kinetic and ITC experiments, as it would eliminate or largely reduce catalytic turn over, while preserving the binding event. The use of the comparatively stable PCP for crystallography may also provide more insight into the binding behaviour of PRPP in ATP-PRT enzymes analogous to results on type I PRT enzymes.

5.4.2 Are there more key interactions in the inhibitory signal communication?

The rational mutagenesis of *Cje*ATP-PRT has led to the discovery of a residue crucial for signal transmission through the enzyme. Thereby a first success has been achieved in unravelling detailed information about the transmission mechanisms. By placing a strong focus on the crystallographic analysis and comparative computational modelling, the involvement of single residues and crucial contacts in the domain movements and interface rearrangements should become clear and allow for further identification and characterisation of these key interactions.

5.4.3 Can an ATP-PRT long form be converted into a fully functional short form?

Preliminary experiments have shown that the truncated *Cje*ATP-PRT Core is not able to interact with separately purified HisZ *in vitro*. While this analysis is far from complete, it opens the opportunity to investigate what truly distinguishes long and short form ATP-PRT chains. Combinations of truncation mutants and HisZ proteins from several different organism sources should provide a deeper insight into essential components of the interaction site between ATP-PRT and HisZ in the hetero-octameric complexes. Another promising approach would be the rational design of *Cje*ATP-PRT Core to be more short form-like, by adjusting or replacing structure motifs or residue ranges with their short form equivalents. The most likely starting points hereby are the helix $\alpha 8$ and the strand $\beta 5$, which form large parts of the interface with HisZ. If successful, this experiment would provide more evidence for a dimeric common ATP-PRT ancestor.

5.4.4 Can the evolutionary recruitment of the ACT domain be simulated?

The catalytic portion of *Cje*ATP-PRT, the Core mutant, on its own has shown to be a relatively ineffective catalyst. By creating an artificial fusion protein with a contemporary P_{II} protein, it might be possible to restore its full catalytic efficiency, thereby re-creating the evolutionary event that led to the contemporary ATP-PRT long form enzymes. Depending on the choice of the fusion partner, this could potentially also replace the allosteric effector. Many common P_{II} proteins, like GlnB or GlnK, are known to bind ATP, which could be disruptive to the ATP-PRT activity. The ligand binding sites of P_{II} protein involve long loops though, which are not present in ATP-PRT ACT domains. Consequently it needs to be carefully considered which protein is used and which modifications are necessary for this fusion experiment.

Chapter 6

Experimental procedures

6.1 General methods

6.1.1 Multiple sequence alignments

Amino acid sequence alignments used in this thesis were generated using either ClustalW2[102] (http://www.ebi.ac.uk/Tools/msa/clustalw2/) or Clustal Omega[84] (http://www.ebi.ac.uk/Tools/msa/clustalo/). For visual output the alignments were modified using ESPript[88] 3.0 (http://espript.ibcp.fr/ESPript/ESPript/).

6.1.2 Phylogenetic trees

To generate the dendrograms displayed in this thesis the distance matrices created by Clustal Omega were visualised with the program FigTree.[85]

6.1.3 Protein structure figures

All figures of protein structures were created using the PyMOL Molecular Graphics System (version 1.7.4.0, Schrödinger, LLC).[99] Shown ligand density

maps represent the unbiased expected ligand density (Fo-Fc) contoured at 3.0 sigma.

6.1.4 Water

Water used in all experiments was purified by a Millipore Milli-Q® system. If required for molecular biology or cell culturing the water was sterilised by autoclaving prior to use.

6.1.5 Metal ion removal

When required, divalent metal ions were removed from solutions by treatment with Chelex® 100 resin, 100–200 mesh (Bio-Rad). After addition of the resin the solution was stirred over night and the resin was removed by filtration (0.22 μm membrane) afterwards. The pH of all Chelex® treated solutions was re-adjusted according to the requirements.

6.1.6 pH measurements

All pH measurements were performed using a Mettler Toledo SevenCompact™ pH/Ion meter S220 (Global Science) with a InLab® Expert Pro-ISM pH probe. The pH of solutions was adjusted by using HCl (1 M or 10 M) or glacial acetic acid, and NaOH (1 M or 10 M).

6.1.7 Antibiotics

1000-fold concentrated antibiotic stocks of ampicillin (100 mg mL^{-1}), Kanamycin (50 mg mL^{-1}), and chloramphenicol (25 mg mL^{-1}, solvent ethanol) were prepared in batches, filter sterilised, and stored at $-80°C$. For any application requiring antibiotics these stocks were thawed and diluted accordingly.

Spectinomycin was used at an effective concentration of 0.1 mg mL^{-1} and always prepared immediately prior to use.

6.1.8 Cell lines

The *E. coli* cell lines used in all experiments were as follows. High-copy number strains for plasmid propagation were either TOP10 (Invitrogen) or Stellar™ (Clontech). For protein expression BL21* (DE3), BL21* (DE3) pLysS, or BL21* (DE3) pBB540 + pBB542 (referred to as Chaperone3) cell lines were used.

6.1.9 Glycerol stocks

Cell stocks of all plasmid-containing strains were prepared by mixing 250 μL of 60% glycerol (sterile) with 750 μL of an overnight culture in a 1.7 mL screw-top tube. Stocks were flash frozen in liquid nitrogen and stored at −80°C.

6.1.10 Chemically competent cells

For the preparation of competent cells a 100 mL LB culture containing the appropriate antibiotics was inoculated from a 5 mL overnight culture of the desired strain and grown until an optical density of just over 0.4 was reached. The cells were then centrifuged at 4000 *g* for 15 minutes in two sterile conical 50 mL centrifuge tubes. After decanting the LB media the cells were resuspended in 25 mL freshly prepared sterile $CaCl_2$ per tube by gentle vortexing and left to incubate on ice for 30 minutes. The cell suspension was centrifuged again at 4000 *g* for 15 minutes and the supernatant decanted. The cells were resuspended in 2.5 mL sterile $CaCl_2$ per tube by gentle vortexing and left on ice for 2–4 hours thereafter. Glycerol was added to a final concentration of 15 % v/v and the suspension was quickly aliquoted (50 or 100 μL) and flash frozen in liquid nitrogen.

All centrifugation steps were performed in pre-chilled centrifuges at 4°C and all equipment used in the process was kept cool at all times.

6.1.11 Transformation

Chemically competent cell aliquots (50 or 100 μL) were slowly thawed on ice prior to the addition of 1–2 μL purified plasmid (50-100 ng μL^{-1}). The cell suspension was gently mixed and kept on ice for 30 minutes prior to heat shock. The tubes were then placed in a warm water bath (42°C) for 1 minute and cooled down on ice for 1 minute afterwards. The cells were regenerated by addition of 250–500 μL super optimal broth with catabolite repression (SOC) medium and incubated at 37°C and 180 rpm for 1 hour. A portion of the cell suspension was then spread on LB-agar plates containing the appropriate antibiotics and incubated overnight at 37°C.

SOC medium: 2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 20 mM glucose. Filter sterilised in aliquots.

6.1.12 Plasmid extraction and purification

Plasmids were purified using one of the following kits: High Pure Plasmid Isolation Kit (Roche), ChargeSwitch® Pro Filter Plasmid Miniprep Kit (Invitrogen), or NucleoSpin® Plasmid EasyPure (Macherey-Nagel). Cells were grown as 5 mL overnight cultures in glass test tubes containing the appropriate antibiotics and harvested by centrifugation. The extraction and purification was performed according to the manufacturer's instructions.

The concentration of the purified plasmid was measured by absorption at 260 nm using an appropriately blanked NanoDrop® ND-1000 spectrophotometer.

6.1.13 Agarose gel electrophoresis

Per gel 0.7–1.0 % (w/v) agarose was dissolved in 30 mL Tris-acetate-

EDTA (TAE) buffer using a microwave and equilibrated in a 60°C oven before the addition of 3 μL SYBR® Safe DNA gel stain. The mixture was then poured into a sealed gel tray including a comb and left to set. The samples were mixed 5:1 with a sample loading dye prior to loading. Typically 20 μL of sample were loaded per well. One well on each gel held 12 μL of DNA ladder (1.2μL of 1 kb Plus DNA ladder (Invitrogen) diluted to 10 μL mixed with 2 μL sample loading dye) to estimate the sample fragment size(s). The electrophoresis was performed for 45 min at 85 V in TAE buffer in a Mini-Sub® Cell GT (Bio-Rad) and the resulting gel was visualised and photographed under UV (302 nm) using a Molecular Imager® Gel Doc™ XR (Bio-Rad).

TAE buffer: 50 mM Tris, 1 mM EDTA, 20 mM acetic acid

Sample loading dye (6x): 60 mM Tris-HCl, 60 mM EDTA, 0.2% (w/v) orange G, 0.05% (w/v) xylene cyanol ff, 60% (v/v) glycerol

Alternatively pre-cast 1.2% E-Gel® or 0.8% E-Gel® CloneWell™ agarose gels (Invitrogen) were loaded and run according to manufacturer's instructions in a E-Gel® iBase™ Safe Imager™ system which also allowed the real-time monitoring of the sample separation and safe band excision via blue light (480 nm).

6.1.14 Polymerase chain reaction (PCR)

All PCR was performed in a Veriti® 96-well Thermal Cycler (Applied Biosystems) or an iCycler (Bio-Rad) using Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific), if not stated otherwise. Primers were synthesised by Invitrogen and dissolved in either autoclaved water or TE Buffer (10 mM Tris/HCl, 0.1 mM EDTA, pH 8.0) to a stock concentration of 100 μM. A 10 μM working stock was prepared prior to use for most applications. All reactions were performed in a 50 μL volume.

6.1.15 Site-directed mutagenesis (SDM)

For mutagenesis *Pfu Ultra* High-Fidelity Polymerase (Agilent Technolo-

gies) was used according to instructions with 30 repeats in the cycling protocol. Complementary mutagenesis primers were designed to a length of approximately 40–50 base pairs with the point of mutation in the centre leaving sufficient overlap (15–20 base pairs) on both sides. The methylated template DNA was digested using the restriction enzyme DpnI supplied by either Life Technologies or New England BioLabs afterwards and the resulting product verified by agarose gel electrophoresis.

6.1.16 Gateway® Cloning

The Invitrogen Gateway® Cloning system makes use of the site-specific recombination properties of bacteriophage λ Integrase and Excisionase together with the *E. coli* Integration Host Factor to facilitate the transfer of heterologous DNA sequences between vectors.[103] It features the two highly specific enzyme mixes BP Clonase™ and LR Clonase™ as well as several recombination vectors (pDONR™, pENTR™, and pDEST™).

All genes cloned by this method were amplified from genomic DNA by PCR in two steps (nested PCR). Whilst the first step amplified the gene itself, specific recombination sites (attB sites) were added in the second step using generic primers. The second step also introduced a N-terminal tobacco etch virus protease cleavage site between the attB1 site and the start-codon. Primers used for nested round one were designed by the same method for all genes. Design scaffold for the first round primers and the generic second round primers used are listed below. XXX symbolises a 15–18 nucleotide long sequence of the beginning or end of the cloned gene including start or stop codon.

1st R nested FWD	5' GGCAGCGGCGCGXXX 3'
1st R nested REV	5' GAAAGCTGGGTGXXX 3'
2nd R nested FWD	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAAACC
continued	TGTATTTTCAGGGCAGCGGCGCG 3'
2nd R nested REV	5' GGGGACCACTTTGTACAAGAAAGCTGGGT 3'

The linear DNA construct is then transferred into pDONR™221 by performing the BP Clonase™ reaction according to the manufacturer's instructions. During this recombination the original insert in pDONR™221 encoding for a *ccd*B gene is replaced by the gene of interest. The resulting

entry clone is transformed into chemically competent TOP10 cells and plated on LB-agar containing kanamycin for selection. As the protein encoded by the *ccd*B gene terminates cell growth it acts as a negative selection and only correctly recombined vectors will produce colonies.

Colonies were picked and grown for plasmid extraction. The purified entry clone is sequence verified before proceeding to the second recombination step using LR Clonase™ mix according to the instructions provided. Here the desired gene is transferred out of the pDONR™221 vector into a destination vector with N-terminal tag (either pDEST™15 or pDEST™17) creating the final expression construct. The reaction mixture is transformed into TOP10 cells and plated on LB-agar containing ampicillin for selection. Again only correctly recombined vectors produce colonies due to the cell growth inhibiting properties of the *ccd*B casette contained in the parental pDEST™ vectors. One or two of the resulting colonies were picked and grown over night to prepare both a glycerol stock and purified expression clone, which was transformed into *E. coli* expression strains and used for heterologous expression thereafter.

6.1.17 Colony PCR

Whenever a newly created construct was transformed into a plasmid propagation strain the presence of the desired gene in the resulting clones was verified by PCR. A fraction of each picked colony from the overnight plate was placed into a PCR reaction as the template. Usually the reaction mixture contained a gene specific forward primer and a plasmid specific reverse primer and amplification was performed using Taq DNA Polymerase (Roche) with an elongated initial denaturation step (5 min). After agarose gel electrophoresis showed the expected construct size the remainder of the colony was used to inoculate an overnight culture (5 mL) for plasmid preparation.

6.1.18 DNA sequencing

All DNA sequencing was performed at Canterbury Sequencing and Genotyping using an Applied Biosystems 3130xl Genetic Analyzer with BigDye Terminator 189 v3.1 (Applied Biosystems) sequencing chemistry. 300 ng of purified double-stranded plasmid and were supplied for each sample while sequencing primers, either M13 (pDONR^{TM}221 vector) or T7 (pDEST and pET expression plasmids), were diluted to 3.2 μM and supplied as 50 μL aliquots.

M13 FWD	5' GTAAAACGACGGCCAG 3'
M13 REV	5' AACAGCTATGACCATG 3'
T7 FWD	5' TAATACGACTCACTATAGGG 3'
T7 REV	5' CTAGTTATTGCTCAGCGGTG 3'

6.1.19 Culture media

Lysogeny broth (LB) media for bacterial growth were prepared by dissolving tryptone (10 g L^{-1}), yeast extract (5 g L^{-1}), and NaCl (5 g L^{-1}) in water and sterilised by autoclaving. Antibiotics were added as required before use.

For solid media LB-agar (Miller's) (37 g L^{-1}) was dissolved in water, sterilised by autoclaving, and stored at room temperature until use. The LB-agar was heated to a boil using a microwave oven until all solid was dissolved. The solution was then allowed to cool down before the addition of the required antibiotics and the plates were cast in circular petri dishes.

6.1.20 Cell culturing and protein expression

Expression strains were cultured in 1, 2, or 3 L baffled conical flasks in an Ecotron shaker/incubator (INFORS HT) or an Innova®44 stackable incubator shaker (New Brunswick Scientific) at 180 rpm. In general not more than 30% of the flask's volume was filled for optimal aeration. The sterilised medium was inoculated with an overnight culture of the desired

strain to an optical density of 0.1 and the cells were left to grow at 37°C until the mid-logarithmic growth phase was reached before isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 to 0.5 mM. After induction the culture was left to grow either at 180 rpm and 37°C for 2–4 hours or at 180 rpm and 23°C overnight before harvest.

All expression vectors (pDEST or pET) used are under the control of a T7 promotor and all expression strains used contain the T7 RNA polymerase gene under the control of a inducible *lac* promotor. IPTG binds to the *lac* repressor and induces the expression of the T7 RNA polymerase which in turn leads to the over-expression of the desired gene from the expression vector. This basic expression system can be further modified. In the case of BL21*(DE3) pLysS the additional plasmid pLysS contains the gene for the T7 lysozyme which acts as a repressor of the basic T7 RNA polymerase activity (in a non-induced culture) allowing the fine-tuning of the protein expression via IPTG. The pBB540 and pBB542 plasmids of the Chaperone3 strain produce the chaperone proteins GroEL/GroES which counter misfolding and aggregation of the over-expressed protein of interest.

6.1.21 Cell harvesting

Liquid cultures were harvested by centrifugation in a Fiberlite™ F9-6 x 1000 LEX Fixed-Angle rotor (Thermo Scientific) in either 1 or 0.8 L bottles, or 0.4 L bottles in the appropriate adaptors at 12000 g for 15 minutes at 4°C. The supernatant liquid was decanted and the pellet was scraped into sterile 50 mL falcon tubes. Small cultures were directly harvested in sterile 50 mL tubes using a Fiberlite™ F14-14 x 50cy Fixed-Angle rotor (Thermo Scientific). Harvested cell pellets were either immediately lysed or frozen and stored at −80°C.

6.1.22 Cell lysis

Cell lysis was performed by one of three different methods. In all three cases the cell pellets were thawed on ice and dissolved in chilled lysis buffer

to a more or less viscous (depending on the methods requirements), homogeneous suspension.

Most commonly cells were lysed by sonication using an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International). The cell suspension was kept in a glass beaker in a tightly packed ice bath. Sonication was performed in multiples (usually three to four repeats) of 5 minutes at 70 % power and 30 % pulsation until the suspension cleared or its viscosity significantly decreased.

Small cultures were typically lysed by chemical detergent using BugBuster® Protein Extraction Reagent (Novagen). 10 mL of 1x BugBuster® solution (in lysis buffer) were used per gram of cell pellet. Benzonase® Nuclease (Novagen) was added (≤ 1 μL) to the mixture and the resuspended cells were incubated at RT for an hour with gentle shaking/rocking.

The third method, cell disruption, made use of a M-110P Microfluidiser® (Microfluidics). The cell pellet was resuspended in a Benzonase® (≤ 1 μL) containing lysis buffer and left for approximately 10 minutes before disruption. The suspension was passed three times through the well-cooled pressure-chamber at 18000 psi and monitored to avoid overheating.

Removal of the cellular debris was achieved by centrifugation at 40000 g and 4°C for 30 minutes in a Fiberlite™ F21-8 x 50y Fixed-Angle Rotor (Thermo Scientific).

6.1.23 Protein purification

All protein constructs were engineered to have either a poly-histidine or a glutathione S-transferase tag. The first step of all purifications therefore was affinity chromatography, followed by the optional removal of the tag using TEV protease and a final size exclusion chromatography step. All chromatography was performed on Bio-Rad Biologic DuoFlow and GE Healthcare ÄKTApurifier™ 10 machines. All buffers used for chromatography were filtered (0.22 μm) and degassed under vacuum.

Samples were pre-filtered using a 0.8 μm filter and then loaded into either a 10 mL, 50 mL, or 150 mL Superloop™ (GE Healthcare) while simultaneously being filtered (0.22 μm). The eluate was collected in 96-well plates (2 mL fractions). Elution from the columns was monitored at 280 nm and 260 nm wavelength, and also 214 nm when possible. Elution peaks were analysed by SDS-PAGE and for ATP-PRT activity the first time any purification step was performed or when major changes where made to the protocol. Verified ATP-PRT fractions were pooled and prepared for the next chromatographic step.

6.1.24 Immobilised metal affinity chromatography

Throughout all IMAC purifications either 5 mL HiTrap™ TALON® crude (Co^{2+}) or 5 mL HiTrap™ HP (Ni^{2+}) columns (GE Healthcare) were used. The column was washed with at least 3 column volumes of water prior to equilibration with 5 column volumes of washing buffer (50 mM phosphate, 500 mM NaCl, 5 mM $MgCl_2$, 20 mM imidazole, pH 8.0). The clarified supernatant was then loaded onto the column and the column was washed with another 5–10 column volumes of washing buffer. A gradient increasing the concentration from 20 to 250 (Co^{2+}) or 500 mM (Ni^{2+}) imidazole over 7 column volumes was applied to elute the bound protein of interest before washing the column with 3 column volumes of elution buffer (50 mM phosphate, 500 mM NaCl, 5 mM $MgCl_2$, 250 or 500 mM imidazole, pH 8.0) and a final 5 column volumes of washing buffer. The flow rate was kept constant at 3 mL min^{-1} throughout the run except for the injection of the protein solution which was held at 1 mL min^{-1}. All steps mentioned above were routinely performed using a pre-programmed method and only changed occasionally. The sodium chloride concentration was altered for some proteins but was kept at a minimum of 300 mM for optimal protein-column interaction. For storage purposes the column was first washed with water to remove all salt and then transferred into 20 % v/v ethanol and kept at 4°C.

During solubility tests disposable His SpinTrap™ columns (GE Healthcare) were used according to the instruction manual to allow for the isola-

tion of the his-tagged protein from the separated soluble and insoluble crude lysate. The protein was then identified by consecutive SDS-PAGE and ATP-PRT activity assays.

The *Cje*ATP-PRT has an intrinsic affinity to the IMAC columns, which means it bound even when it did not contain a poly-histidine tag and eluted at approximately 100 mM (Co^{2+}) or 200 mM (Ni^{2+}) imidazole. It was therefore unfavourable to use IMAC for purification other then the initial separation.

6.1.25 Glutathione sulfur-transferase affinity chromatography

GST affinity chromatography was performed in a very similar fashion to the IMAC. The columns used for this purification step were 5 mL GSTrap™ HP columns (GE Healthcare) and column handling and protocol were identical to IMAC with the buffers being the only difference. Wash (and lysis) buffer in this case was 1x phosphate buffered saline (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and the elution buffer was 50 mM Tris/HCl pH 8.0 with 10 mM reduced glutathione.

Besides the first step of purification GST affinity chromatography was also used later on as reverse purification step to remove the GST tag, especially in the case of the *Cje*ATP-PRT Core mutant. When used as a downstream purification step buffer exchange of the protein solution into 1x phosphate buffered saline was performed directly before the run.

6.1.26 Size exclusion chromatography

Fractions containing the desired protein were pooled and concentrated to about 2–3 mg mL^{-1} if not already higher. SEC was performed using a HiLoad™ 26/60 Superdex™ 200 prep grade or a HiLoad™ 16/60 Superdex™ 75 prep grade column (GE Healthcare), at 4°C in 10 (or 50) mM Tris/HCl, 100 (or 150) mM NaCl, 5 mM $MgCl_2$, pH 8.0. The elution flow rate was usually kept at 1–1.2 mL min^{-1}. All fractions containing the pure protein of interest were pooled and concentrated afterwards.

For the purpose of molecular weight determination a Superdex™ 200 10/300 GL column (GE Healthcare) was used. The column was equilibrated and then calibrated by running through standard proteins of known sizes in the range of 29 to 669 kDa and Blue Dextran (Gel Filtration Calibration Kit HMW, (GE Healthcare), Molecular Weight Marker Kit MW-GF-1000 (Sigma)) in above stated buffer at a flow rate of 0.2 mL min^{-1}. The sample was analysed using the same buffer and flow rate. Samples and standards (usually at a concentration of 1 mg mL^{-1}) were injected in 500 μL volume using a static loop (either 0.5 or 1 mL volume). Elution times were normalised by the void volume (Blue Dextran).

6.1.27 Tobacco etch virus protease treatment

Frozen stocks of TEV protease were thawed and added directly to the tagged protein solution so that the end concentration of TEV protease was $\frac{1}{10}$ to $\frac{1}{100}$ of that of the tagged protein. The mixture was then incubated for 1 hour at 37°C and 180 rpm and either used immediately in the next purification step or left at 4°C over night before advancing in the purification. In the case of the *Cje*ATP-PRT wild type enzyme this 4°C incubation had to be prolonged to 2–3 days for optimal cleavage of the tag.

6.1.28 Protein concentration and buffer exchange

Protein solutions were concentrated using 10000 Da (or 5000 Da) molecular weight cut off (MWCO) devices Vivaspin® Turbo 15, Vivaspin® 20, or Vivaspin® 500 (Sartorius). 5000 Da MWCO devices were only used for *Cje*ATP-PRT ACT preparations. All filters were rinsed with water prior to use. Buffer exchange was achieved by repeated concentration and dilution cycles, using the desired buffer, in these filtration units.

Large quantities of concentrated protein solution were buffer exchanged by using either a HiPrep™ 26/10 Desalting column (GE Healthcare) or a 50 mL Bio-Scale™ Mini Bio-Gel® P-6 desalting cartridge, equilibrated with at least 3 column volumes of the desired buffer. The protein sample was then

loaded and eluted by applying a continuous flow rate. Any salt and buffer remaining on the column were removed afterwards by thoroughly washing the column with water (2–3 column volumes) before running it into 20 % ethanol (2 column volumes) for storage.

6.1.29 Determination of protein concentration

For all purposes protein concentrations were measured by absorption at 280 nm with an appropriately blanked Nanodrop® ND-1000 spectrophotometer using 1.5 μL of purified protein solution for each measurement. The concentration was calculated from the absorption using the molar extinction coefficient values calculated from protein sequences using the ProtParam tool on the ExPASy Proteomics Server. Used extinction coefficients were:

*Cje*ATP-PRT wild type and single site mutants	14440 M^{-1} cm^{-1}
*Cje*ATP-PRT Core	7450 M^{-1} cm^{-1}
*Cje*ATP-PRT ACT	6990 M^{-1} cm^{-1}
*Mtu*ATP-PRT wild type	19940 M^{-1} cm^{-1}
*Eco*PPase	22920 M^{-1} cm^{-1}

6.1.30 Protein storage

All enzymes were generally kept at 4°C if they were intended for immediate (up to 3 days) use to decrease losses due to the freeze and thaw processes. For long term storage proteins were concentrated to 5–20 mg mL^{-1}, aliquoted in either 50 or 100 μL aliquots in 0.6 mL micro-centrifuge tubes, flash-frozen in liquid nitrogen, and kept at −80°C. Frozen samples were rapidly thawed before use.

6.1.31 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SDS-PAGE was performed using either NuPAGE® 10% Bis-Tris Gel 1.0 mm 12-well (Invitrogen) pre-cast gels in NuPAGE® MOPS SDS Running

Buffer (Invitrogen) or Bolt® 10% Bis-Tris Gel 1.0 mm 12-well (Life Technologies) pre-cast gels in Bolt® MES SDS Running Buffer (Life Technologies) or Bolt® MOPS SDS Running Buffer (Life Technologies). Electrophoresis was performed using a XCell SureLock™ Electrophoresis Cell (Invitrogen) at 200 V for 50 min for NuPAGE® gels and a Bolt® Mini Gel Tank (Life Technologies) at 165 V for 35 min (MES) or 45 min (MOPS) when Bolt® gels were used. Samples were mixed with 4x NuPAGE® LDS Sample Buffer or 4x Bolt® LDS Sample Buffer (Life Technologies) respectively. To allow for complete denaturation samples were mixed with dithiothreitol (DTT) and boiled at 95°C for 2–3 minutes prior to loading. To estimate protein size Novex® Sharp Pre-Stained Protein Standards (Invitrogen) was run in one lane of each gel.

6.1.32 Visualisation of PAGE gels

All PAGE gels were stained in square petri dishes with hot stain solution (0.1% (w/v) Coomassie brilliant blue R-250, 10% (v/v) glacial acetic acid, 40% (v/v) methanol) for a minimum of 5 minutes before being destained with hot destaining solution (10% (v/v) glacial acetic acid, 40% (v/v) methanol) for another 5 minutes to remove excess dye. Gel and solution were heated each time in a microwave oven until almost boiling and rocked gently during the incubation time. Best resolution was achieved by destaining the gel afterwards in water over night with gentle rocking. Photographs of the destained gels were taken on a white light transilluminator in a Molecular Imager® Gel Doc™ XR (Bio-Rad).

6.1.33 Standard kinetic assay

The standard assay for ATP-PRT activity monitored the production of PR-ATP directly by gain of absorbance at 290 nm, based on the method of Morton and Parsons.[28] All kinetic measurements were made using a Varian Cary 100 UV-visible spectrophotometer. All reactions were measured in stoppered 1 cm pathlength quartz cuvettes in a total assay volume of 1 mL. Standard assays were carried out at 25°C and all assay solutions were

equilibrated to temperature and baseline absorbance before initialisation with PRPP, or occasionally ATP-PRT.

Standard reaction mixtures contained variable concentrations of PRPP and ATP in 50 mM Tris/HCl, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 8.5 as well as the enzymes ATP-PRT and *Eco*PPase, with the *Eco*PPase always being in at least 5-fold excess. Assay buffers and all assay stock solutions were prepared with Chelex-treated water to remove divalent metal ions. ATP stock solutions were mixed 1:1 with $MgCl_2$ in order to saturate the substrate with Mg^{2+} without significantly altering the free Mg^{2+} concentration of the assay buffer.

Initial rates of reaction were measured as a least-squares fit of the initial rate data using Cary WinUV Kinetics Application (version 3.00, Varian). A unit of enzyme activity was defined as the production of 1 μmol min^{-1} of PR-ATP ($\varepsilon = 3.6 \cdot 10^3$ M^{-1} cm^{-1}) at 25°C, and was calculated from the measured gain of absorbance at 290 nm using Beer's Law. The values of the kinetic parameters were determined by fitting data to the Michaelis-Menten equation using the software GraFit (version 5.0.13, Erithacus Software Limited).

6.1.34 Determination of ATP and PRPP concentration

The concentrations of the substrate solutions (ATP and PRPP) were determined each time before other kinetic measurements were performed using the standard assay for ATP-PRT activity. The concentration of the substrate not being measured was used in excess (usually 20x K_M) and the substrate for which the concentration was being measured was used at a limiting concentration. The secondary substrate saturation was checked by addition of more of the limited substrate after the reaction had reached completion. The change in absorbance, between that before initiation of the reaction and that at completion, was measured as ΔA_1. A control reaction, absent in enzyme, was used to measure the change in absorbance due to the addition of PRPP for reaction initiation (ΔA_2). The total change in ab-

sorbance ($\Delta A_1 + \Delta A_2$) was converted to the concentration of the limiting substrate in the cuvette using Beer's Law. From this the stock substrate concentration was calculated. All concentration measurements were performed in duplicates or triplicates.

6.1.35 Mass spectrometry

Protein molecular mass was determined by electrospray ionisation using a Micromass LCT Classic, Bruker maXis 3G. Highly concentrated protein samples were diluted with water to a final concentration of 1 mg mL^{-1} to minimise the effect of buffer components during the measurements. Alternatively the protein samples were buffer exchanged into 5 mM ammonium bicarbonate.

6.1.36 Circular dichroism

All CD experiments were performed using a JASCO J-815 Spectropolarimeter. CD spectra were recorded from 250 nm to 190 nm at 25°C using a 1 nm data pitch, 1 s response, and 1 nm bandwidth. 2 mL of protein solution (0.03 mg ml^{-1}) in a 3 mL quartz cuvette with 1 cm path length were used for each measurement. All measurements were taken in water and blank spectra of water were recorded before each experiment. The buffer-subtracted data were smoothed by performing a five-point moving average.

6.1.37 Differential scanning fluorimetry

Melting temperatures in the presence of different water soluble additives were determined by DSF using an iCycler iQ5 Multicolour Real-Time PCR Detection System (Bio-Rad). The method used was based on that of Nordlund et al.[104]

For each condition 100 μL of buffer (50mM Tris/HCl, 150 mM NaCl, pH 8.5) including the desired additive(s) was prepared, to which 5 μL of 250x SYPRO® Orange dye was added. One fifth (21 μL) was mixed with 4 μL of water, representing the blank. To the remaining 84 μL of the condition

solution 16 μL of protein (0.6 mg mL^{-1}) were added and 25 μL each were dispensed into three replicate wells followed by the blank in a fourth adjacent well of a 96-well microplate (vide infra). After all samples and blanks were loaded the plate was sealed with a self-adhesive transparent seal.

The plate was set up and checked for initial protein denaturation which was commonly seen for all *Cje*ATP-PRT samples. The melt proceeded in 0.2°C increments from 20 to 95 °C, with a 20 s dwell time after each temperature rise. Measurements of the fluorescence were made at the end of each dwell time. An Excel spreadsheet with custom VBA-scripted macro was used for data analysis to toggle and dynamically display melt and derivative curves for each sample. The melt temperatures were calculated as the temperature of maximum inflection of the melting curve after subtracting the reading of the blank well. Additives used were: L-histidine (1 mM), ATP (2 mM), AMP (2 mM), and $MgCl_2$ (0, 0.5, and 5 mM)

6.1.38 Differential scanning calorimetry

For DSC experiments protein samples (0.3 to 1.0 mg mL^{-1}) were buffer exchanged into 50 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.5 containing either no ligand, 1 mM histidine, 2 mM AMP, or the mixture of the latter two. All prepared solutions were degassed prior to the experiments. Samples were analysed against their identical buffer using a Nano DSC differential scanning calorimeter (TA Instruments). The sample and reference chamber were thoroughly cleaned with detergent (Decon 90) and water before the chambers were washed with buffer. Cells were loaded by gently pipetting the solutions in, and up and down to remove any residual air bubbles.

The temperature was increased from 20 to 100°C in increments of 1°C per minute under a constant pressure of three atmospheres. Data analysis was performed with the NanoAnalyze™ software version X.1.2 (TA Instruments). The baseline corrected data was fitted using the provided TwoStateScaled model if not stated otherwise.

6.1.39 Isothermal titration calorimetry

Dissociation constants were determined using a VP-ITC microcalorimeter (MicroCal, GE Healthcare) operating at 298 K. All protein solutions (50 to 100 μM) were buffer exchanged into and ligands were made up in binding buffer (10 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.5). Solutions were filtered and degassed under vacuum and the protein concentration was measured by UV absorption prior to the experiments. The syringe and cell were washed multiple times with binding buffer and degassed binding buffer.

For all experiments the cell was loaded with the protein and the syringe contained the ligand (titrant). Titrations were performed with 56 injections (one 2 μL injection followed by fifty-five 5 μL injections) or 29 injections (one 2 μL injection followed by twenty-eight 5 μL injections) of ligand. Heat of dilution experiments were measured independently and subtracted from the integrated data before curve-fitting in Origin (version 7.0, OriginLab®). The initial point (2 μL injection) was routinely deleted to allow for diffusion of ligand across the needle tip during the equilibration period and the used binding model was chosen according to the experiment performed.

6.1.40 Static light scattering

Static light scattering was used to determine the molecular weight and hydrodynamic radius of proteins in solution. Measurements were performed using a Superdex™ 200 10/300 GL column directly coupled to a Viscotek TDAmax multidetector system (Malvern) at constant flow rate of 0.5 mL min^{-1}. Typically only UV/Vis, Refractive Index (RI) and Right angle light scattering (RALS) detections were used.

Before every experiment the system was equilibrated with one column volume of running buffer (10 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.5). For each sample a total amount of 100 μL of protein solution at 1 mg mL^{-1} was injected and run through with one column volume of running buffer. A bovine serum albumin (BSA) standard was run before and after every set of samples and used to calibrate the OmniSEC software (Malvern)

for analysis of the recorded sample peaks.

6.1.41 Analytical ultracentrifugation

Sedimentation velocity experiments were performed in a ProteomeLab™ XL-I analytical ultracentrifuge equipped with UV–visible scanning optics (Beckman Coulter). Protein samples were prepared by dialysis in the reference buffer (10 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.5) over night and then diluted, in sets of three, to different concentrations ranging from 1.2 mg mL^{-1} down to 0.025 mg mL^{-1} using the identical reference buffer in which the dialysis had been performed. High range: 1.2 mg mL^{-1}, 0.4 mg mL^{-1} and 0.13 mg mL^{-1}. Low range: 0.1 mg mL^{-1}, 0.05 mg mL^{-1} and 0.025 mg mL^{-1}.

Reference buffer (400 μL) and samples (380 μL) were loaded into 12-mm double-sector cells with quartz windows, and the cells were then mounted in an An-50 Ti eight-hole rotor. Proteins were centrifuged at 50000 rpm and 20°C, and radial absorbance data were collected at 230 nm wavelengths, due to the low absorbance of *Cje*ATP-PRT at 280 nm, in continuous mode every 8 min without averaging.

AUC data was fitted to a continuous distribution c(s) model using SEDFIT.[90,105] The partial specific volume ($\overline{v}$) of the proteins, buffer density (0.7482 g mL^{-1} for the *Cje*ATP-PRT wild type, 0.7434 g mL^{-1} for the *Cje*ATP-PRT Core) and buffer viscosity (1.00618 cp) were calculated using the program SEDNTERP.[106] The ratios of the frictional coefficients ($\frac{f}{f_0}$) calculated by SEDFIT from fitting the c(s) models were 1.27 for the *Cje*ATP-PRT Core and 1.29 for the *Cje*ATP-PRT wild type.

6.1.42 Small angle X-ray scattering measurements

Measurements were performed at the Australian Synchrotron SAXS/ WAXS beamline equipped with a Pilatus detector (1 M, 170 mm x 170 mm, effective pixel size, 172 x 172 μm). The wavelength of the X-rays was 1.0332 Å. The sample-detector distance was 1.6 m, which provided an s range of

0.0126 to 0.6014Å^{-1}. s is defined as the magnitude of the scattering vector, which is related to the scattering angle (2Θ) and the wavelength (λ) as follows: $s = (\frac{4\pi}{\lambda})sin\Theta$.

Protein samples ($\sim$ 10 mg mL^{-1}) filtered (0.22 μm) prior to use to avoid aggregation. For the capillary set-up of the beam line samples were buffer exchanged and diluted (to 0.5 mg mL^{-1}, 1 mg mL^{-1}, and 2 mg mL^{-1}) using the appropriate buffer, transferred into 96-well plates, and degassed by centrifugation (10 min at 4000 g). Scattering data was collected either following elution from a SEC column (Superdex$^{\text{TM}}$ 200 5/150), pre-equilibrated (50 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0) with or without effectors PRPP (1 mM), AMP (2 mM), or L-histidine (1 mM) or directly from a sealed 96-well plate containing the desired conditions.

Data were collected from a 1.5 mm glass capillary at 27°C at 2 s intervals. Two-dimensional intensity plots from the top of the size-exclusion chromatography peak (column set up) or the sample conditions (plate set up) were radially averaged, normalised to sample transmission and background subtracted using ScatterBrain (www.synchrotron.org.au/index.php/aussync beamlines/saxswaxs/software-saxswaxs).

6.1.43 Small angle X-ray scattering data analysis

Scattered intensity (I) was plotted versus s. All samples were checked for non-linearity in the increase of intensity at low s (indicative of aggregation). Guinier plots were linear for $s \cdot R_g < 1.3$ as calculated by PRIMUS.[107] Indirect Fourier transform was performed using GNOM[95] to provide the pair wise distribution function P(r), which determines both the relative probabilities of distances between scattering centres and the maximum dimension of the scattering particle (D_{max}). Theoretical scattering curves were generated from atomic coordinates and compared with experimental scattering curves using CRYSOL.[96] To assess and compare the quality of the fits in CRYSOL χ values for the fits were compared.

6.1.44 Crystallisation

Initial screening was performed using a mosquito® Crystal robot (TTP Labtech), the PACT *premier*™ HT-96, JCSG-*plus*™ HT-96, Clear Strategy™ I HT-96 and Clear Strategy™ II HT-96 crystallisation screens (Molecular Dimensions), and 96-well sitting drop iQ plates (TTP Labtech). For each condition drops of 400 nL concentrated pure protein and 400 nL of reservoir solution were mixed and placed next to the 40 μL reservoir by the robot. Plates were sealed, incubated at 20°C, and checked regularly for crystal formation. After a suitable condition had been identified the crystallisation was scaled up and fine screened in 24-well VDX plates (Hampton Research) at 20°C using hanging-drop vapour diffusion. Protein solutions were mixed 1:1 (v/v) with the reservoir solution. Here the drop size was 2 μL or 4 μL and the reservoir solution volume 500 μL. Typically crystals were fully formed after 1–3 days.

Co-crystallisation was achieved by mixing the desired ligand with the crystallisation condition prior to mixing with the protein solution. This generally required further screening either manually or using the mosquito® Crystal robot. In some cases the ligand was soaked into previously formed crystals. This was achieved by the addition of the ligand (dissolved or diluted in the condition) to the drop containing the crystal and incubation for a set amount of time. Crystals for collection were mounted on loops, transferred briefly into the appropriate cryoprotectant, flash frozen in liquid nitrogen and stored in Australian Synchrotron Pucks for transportation.

6.1.45 X-ray data collection and structure determination

All X-ray diffraction data was collected at the Australian Synchrotron (AS). Pre-mounted crystals stored in AS Pucks were transported in a dry shipper (Taylor-Wharton CX100) and transported as checked-in luggage. Data collection at the Australian Synchrotron used either the Macromolecular Crystallography (MX1) or Micro Crystallography (MX2) beamlines[108]

and were processed using X-ray Detector Software (XDS),[109] Aimless and Truncate[110] (CCP4 suite[111]).

For novel structures initial phase estimates were calculated by molecular replacement with MolRep[112] or Phaser[113] (CCP4 suite) using the available *Mtu*ATP-PRT and *Eco*ATP-PRT structures (PDB codes: 1NH7, 1H3D). The R_{free} set (5 %) was chosen automatically. Structures of *Cje*ATP-PRT Core were solved by molecular replacement, using the catalytic domains of wild-type *Cje*ATP-PRT (PDB accession code 4YB5) as starting model with all non-protein molecules removed. 10 rounds of rigid-body refinement were performed using Refmac5[114] and the same set of reflections for calculation of R_{free} was carried through.

Structural refinement was performed using RefMac5 with iterative model building using Coot.[115] NCS was included in the initial rounds of refinement and waters were gradually added in the later rounds. TLS refinement was included in the later rounds of refinement. Structures were validated and adjusted using the validation tools of Coot and the MolProbity[116] as well as the wwPDB deposition servers before deposition.

6.2 Methods for Chapter 2

6.2.1 Cloning of *Cje*ATP-PRT

Cloning proceeded as described (section 6.1.16) with a minor change in the primer design strategy. To avoid the relatively long generic primers of the second nested PCR a set of three forward and two reverse primers was used instead.

cjeAPRT0 FWD	5' GGAATGCAAGAAAACACTCGTTTACGC 3'
cjeAPRT1 FWD	5' GGCTTCGAAAACCTGTATTTTCAAGGAATGCAAGAAAAC 3'
2R GERD	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAAAC 3'
cjeAPRT0 REV	5' GGGTGTCACTTTAACATTTTTTCTATAG 3'
cjeAPRT1 REV	5' GGGGACCACTTTGTACAAGAAAGCTGGGTGTCAC 3'

These were used in three consecutive PCR steps as follows:

Step 1: cjeAPRT0 FWD and cjeAPRT0 REV
Step 2: cjeAPRT1 FWD and cjeAPRT1 REV
Step 3: 2R GERD and cjeAPRT1 REV

This resulting construct differed to the one that would have been created by the standard method only by the absence of a short linker sequence (AGCGGCGCG) between the TEV cleavage site and the start codon.

6.2.2 Cloning and purification of *E. coli* PPase

Cloning of the *E. coli ppa* gene (locus names: b4226, JW4185) followed the general Gateway® Cloning procedure using the following primers in the first round of nested PCR:

Gerd Ecoppa FWD 5' GGCAGCGGCGCGATGAGCTTACTCAACGTCCCT 3'
Gerd Ecoppa REV 5' GAAAGCTGGGTGTTATTTATTCTTTGCGCGCTCG 3'

The pDEST™ 17 construct was cloned into the expression strain BL21* (DE3) and usually expressed overnight at 23°C and 180 rpm after induction with 0.5 mM IPTG.

In most cases *Eco*PPase was found to be exclusively in the insoluble fraction after lysis. A basic resolubilisation was performed by dissolving the entire pellet in lysis buffer containing 8 M urea. The solution was centrifuged again at 40000 *g* and 4°C for 30 minutes in a Fiberlite™ F21-8 x 50y Fixed-Angle Rotor (Thermo Scientific). The resulting supernatant contained more than 95 % of the *Eco*PPase and was diluted 10 fold with lysis buffer to reduce the urea concentration. It was then filtered to remove remaining insoluble particles and DNA.

The protein was purified in a one step protocol using IMAC (5 mL HiTrap™ HP column). Resolubilised *Eco*PPase always eluted in two peaks, a first peak around 250 mM imidazole and a second peak at 500 mM imidazole. All fractions of the second peak precipitated immediately after the

column passage and were discarded. Fractions of the first peak were combined, desalted, concentrated, aliqoted, flash frozen in liquid nitrogen and stored at −80°C until use.

The activity of the resulting enzyme was tested in standard ATP-PRT assay conditions. When added after the ATP-PRT reaction had come to the equilibrium point (first absorbance rise) a second rise in absorbance was observed. The total absorbance change was identical to the change measured in an equivalent reaction where the *Eco*PPase had been added prior to start.

6.2.3 Kinetic measurements

Deviations from the described standard assay (section 6.1.33) occurred only during adjustment of the Mg^{2+} and salt concentration. In this case a metal free buffer was used (50 mM Tris/HCl, pH 8.5) and the salts (metal chlorides) were added from Chelex-treated stock solutions at the desired concentration for each measurement. The Mg^{2+} concentration was held constant at 10 mM while determining the Na^+ and K^+ dependence. While determining the Mg^{2+} dependence Na^+ and K^+ concentrations were held constant at 100 mM.

For determination of the kinetic parameters and the binding mechanism initial rate measurements were taken in a grid of PRPP (0.001 mM, 0.0025 mM, 0.005 mM, 0.0075 mM, 0.01 mM, 0.025 mM, 0.05 mM, 0.075 mM, 0.1 mM and 0.2 mM) against ATP (0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM and 2.5 mM) concentrations. The protein concentration used in the cuvette was 50 nM. Results were plotted using Microsoft Excel and Graph Pad Prism.

Similar grids of concentrations (listed below) were used for the inhibition studies with the second substrate at saturating concentrations and a constant enzyme concentration of 267 nM. The resulting data was plotted and analysed by global fitting using pre-programmed inhibition type templates in GraFit (version 5.0.13, Erithacus Software Limited).

K_{iATP}(AMP)	ATP (0.08 mM, 0.2 mM, 0.4 mM, 0.8 mM)
	AMP (0 mM, 0.5 mM, 1 mM, 2 mM)

K_{iPRPP}(AMP)	PRPP (0.03 mM, 0.06 mM, 0.09 mM, 0.12 mM)
	AMP (0 mM, 1 mM, 2 mM, 3 mM)
K_{iATP}(his)	ATP (0.1 mM, 0.2 mM, 0.4 mM, 1 mM)
	histidine (0 mM, 0.02 mM, 0.04 mM, 0.07 mM)
K_{iPRPP}(his)	PRPP(0.045 mM, 0.09 mM, 0.135 mM, 0.18 mM, 0.3 mM)
	histidine (0 mM, 0.01 mM, 0.02 mM, 0.03 mM)

6.2.4 Ligand binding experiments

All ITC experiments were performed as described in section 6.1.39 with the following concentrations used:

Ligand	Enzyme concentration	Secondary ligand in cell
4 mM ATP	150 μM	-
3 mM PRPP	150 μM	-
3.5 mM AMP	100 μM	-
1.5 mM histidine	100 μM	-
10 mM AMP	100 μM	1 mM ATP
4 mM ATP	150 μM	2 mM histidine
2 mM AMP	100 μM	1 mM histidine
1 mM histidine	100 μM	1 mM AMP

AMP and histidine titrations were performed with 56 5 μL injections, while 28 10 μL injections were used for ATP and PRPP titrations.

All data was fitted with a single site model except the histidine titrations which were done using the sequential site model with either 5 sites (histidine alone) or 2 sites (histidine in the presence of AMP).

6.2.5 Crystallisation conditions

Initial crystal screens were performed as described above (section 6.1.44). The usually used protein concentration ranged between 2 and 5 mg mL^{-1} for all crystallisation trials. Three screen conditions were identified that produced crystals of desired quality. All three conditions were scaled up and fine screened for optimal crystal growth. The resulting mixtures are listed below. Histidine co-crystals were put down as 2 μL drops and the cryoprotectant was composed of 20 % v/v PEG 400 and 80 % reservoir solution. For ATP co-crystals the total drop size was 4 μL and 40 % glycerol was added slowly in a 1:1 ratio to the whole drop prior to crystal-looping for cryoprotection.

Screen condition	Ligand	Refined condition
PACT *premier*™ HT-96 F4	1 mM histidine	0.1 M BTP pH 6.5 0.2 M KSCN 13–14 % w/v PEG 3350
Clear Strategy™ I HT-96 E9	5 mM AMP 1 mM histidine	0.1 M Tris pH 7.5 0.1 M $MgCl_2$ 13–15 % w/v PEG 4000
Clear Strategy™ I HT-96 A9	10 mM ATP	0.1 M Na acetate pH 5.5 0.1 M $MgCl_2$ 13–15 % w/v PEG 4000

6.2.6 *Mtu*ATP-PRT

The cloning of *Mtu*ATP-PRT made use of the In-Fusion® HD Cloning Kit (Clontech® Laboratories Inc.) which is designed for the fast and direct cloning of genes into any vector. It was chosen to insert the *hisG* gene from the *M. tuberculosis* strain H37Rv (gene locus: Rv2121c) into pET24a(+) between the NdeI and XhoI restriction sites. For this purpose primers were

designed according to the user manual instructions to generate 5' 15 bp extensions on the *Mtu hisG* for the purpose of the In-Fusion® recombination. The forward primer also included an additional short insert coding for a non-removable 6x histidine tag and the original start codon was mutated into CTG to eliminate the possibility of a later transcription start.

MtuHisGin24aFWD	5' AAGGAGATATACATATGCATCATCATCATCATCATA ATGGTCTGCTGCGGGTCGCGGTTCC 3'
MtuHisGin24aREV	5' GGTGGTGGTGCTCGAGTCATCAGAATCGGCAGAA CCTGAT 3'

The gene was amplified in 30 cycles using Phusion® High-Fidelity DNA Polymerase with an annealing temperature of 50 °C. Both the linear PCR product and the pET24(+) vector were then digested at 37 °C for 1 h with NdeI and XhoI (New England BioLabs® Inc.) and cleaned using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) before performing the In-Fusion reaction according to the manufacturer's instructions. The product was then transformed into Stellar™ competent cells. The plasmid was extracted from a single cell colony and sequence verified before transforming it into the BL21* (DE3) pLysS expression strain.

The construct was expressed for 4 h at 37°C and 180 rpm after induction with 0.5 mM IPTG and harvested normally. The frozen cell pellet was resuspended in lysis buffer (50 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0) and spiked with Benzonase® Nuclease (Novagen) and lysed by sonication (3 x 5 min).

The supernatant contained the majority of the overexpressed *Mtu*ATP-PRT. Purification was performed by IMAC (5 mL TALON® crude column) applying a gradient from 5 mM imidazole to 250 mM imidazole over 35 mL. The wash buffer (50 mM Tris/HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 8.0) only contained 5 mM imidazole since the initial binding test had shown a large portion of the enzyme in the flow through when 20 mM imidazole was used. The second purification step was a SEC run using the combined elution

peak fractions from the IMAC.

Kinetics were performed using the standard assay for *Cje*ATP-PRT in 50 mM Tris/HCl, 100 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, pH 8.5 as well as 50 mM phosphate, 500 mM NaCl, 10 mM $MgCl_2$, pH 8.5 and 20 mM triethanolamine, 200 mM KCl, 10 mM $MgCl_2$, pH 8.5.

Crystals were grown in the condition described by Cho et al.[35] 7–9 mg mL^{-1} highly pure *Mtu*ATP-PRT was mixed 1:1 with 0.1 M MES pH 6.5, 1.8 M $MgSO_4$. Crystals grew over night and were soaked by adding 1 μL of 10 mM ATP to the crystal drop 30 minutes prior to flash freezing. A mixture of 50 % saturated sucrose and 50 % reservoir served as cryoprotectant.

6.3 Methods for Chapter 3

6.3.1 Cloning of the *Cje*ATP-PRT Core and *Cje*ATP-PRT ACT mutants

The truncation mutants of *Cje*ATP-PRT were created using the same Gateway® Cloning protocol as described for the wild type (sections 6.1.16 and 6.2.1) using the following primers.

Name	**Sequence** (5' - 3')
cjeAPRT0 FWD	GGAATGCAAGAAAACACTCGTTTACGC
cjeAPRT0 REV	GGGTGTCACTTTAACATTTTTTCTATAG
cjeAPRT1 FWD	GGCTTCGAAAACCTGTATTTTCAAGGAATGCAAGAAAAC
cjeAPRT1 REV	GGGGACCACTTTGTACAAGAAAGCTGGGTGTCAC
HisG ACT FWD	GGCAGCGGCGCGTCTAAATACATTATGC
HisG ACT2 FWD	GCTTCGAAGGAGATAGAACCATGTCTAAATACATTATGC
cjeH trunc REV	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTCTC TTGCTTGCATTAC
2R GERD	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAAAC
2nd R nested FWD	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAAACC TGTATTTTCAGGGCAGCGGCGCG
2nd R nested2 FWD	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAG ATAGAACCATG

These were used in three consecutive PCR steps for each construct as follows:

*Cje*ATP-PRT Core

Step 1: cjeAPRT0 FWD and cjeAPRT0 REV
Step 2: cjeAPRT1 FWD and cjeH trunc REV
Step 3: 2R GERD and cjeH trunc REV

*Cje*ATP-PRT ACT

Step 1: cjeAPRT0 FWD and cjeAPRT0 REV
Step 2: HisG ACT FWD and cjeAPRT0 REV
Step 3: 2nd R nested FWD and cjeAPRT1 REV

The *Cje*ATP-PRT ACT was also amplified alternatively without a tag for the consecutive cloning into pDEST™14:

Step 1: cjeAPRT0 FWD and cjeAPRT0 REV
Step 2: HisG ACT2 FWD and cjeAPRT0 REV
Step 3: 2nd R nested2 FWD and cjeAPRT1 REV

The resulting genes were sequence verified in pDONR™221 and transferred into pDEST™15 and pDEST™17 expression plasmids via LR-reaction, and into pDEST™14 in case of the tagless *Cje*ATP-PRT ACT. All constructs were transformed into BL21* (DE3) and Chaperone3 cell lines.

6.3.2 Expression conditions

Large scale expressions were performed using the BL21* (DE3) clones containing the pDEST™15 constructs of *Cje*ATP-PRT Core and ACT mutants. Cells were grown to the exponential phase at 37 °C and 180 rpm before the addition of 0.5 mM IPTG for induction. Cells were left to grow for an ad-

ditional 4 hours at 37 °C and 180 rpm before harvest. Lysis was performed according to general methods (section 6.1.22).

6.3.3 Kinetic measurements

Apparent Michaelis-Menten constants and turn over rates for *Cje*ATP-PRT Core were measured under standard assay conditions (section 6.1.33) altering the PRPP concentration (0.01 mM, 0.02 mM, 0.03 mM, 0.05 mM, 0.075 mM, 0.1 mM, 0.15 mM and 0.2 mM) and ATP concentration (0.02 mM, 0.04 mM, 0.06 mM, 0.1 mM, 0.2 mM, 0.4 mM, 1 mM and 2 mM) separately. The used enzyme concentration in the cuvette was consistently 15.8 μM.

The dose response curve to AMP was measured under standard assay conditions with AMP concentrations in the cuvette being 0 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, 2 mM and 5 mM for the Core mutant and 0 mM, 0.2 mM, 0.4 mM, 0.6 mM, 1 mM, 2 mM, 3 mM, 5 mM and 8 mM in case of the wild type.

The dose response curve to PR-ATP was also measured under standard assay conditions with PR-ATP concentrations in the cuvette being 0 mM, 0.002 mM, 0.005 mM, 0.01 mM, 0.02 mM, 0.05 mM and 0.1 mM.

6.3.4 ITC

AMP and ATP binding experiments on the *Cje*ATP-PRT Core mutant were performed as described in section 6.1.39 using an enzyme concentration of 100 μM in the cell and 2 mM nucleotide as titrant.

Histidine binding experiments on the *Cje*ATP-PRT ACT mutant were performed as described in section 6.1.39 using an enzyme concentration of 57 μM in the cell and 1.5 mM histidine as titrant.

Activity measurements of *Cje*ATP-PRT Core in the presence of *Cje*ATP-PRT ACT were performed under the standard assay conditions (section

6.1.33) with constant concentrations of ATP (2 mM), PRPP (0.12 mM), *Cje*ATP-PRT Core (1.7 μM) and histidine (0.5 mM). The concentrations of *Cje*ATP-PRT ACT used were 1.7 μM, 17 μM and 35 μM according to the ratios 1:1, 1:10 and 1:20.

6.3.5 Crystallisation conditions

All procedures were performed as described in section 6.1.44. The *Cje*ATP-PRT Core mutant was crystallised in the following condition with exceptional consistency and reproducibility. It is derived from the PACT *premier*™ HT-96 condition A12 in which the enzyme was found to crystallise during initial screens. The used protein concentration ranged between 5 and 11 mg mL $^{-1}$ in the drop with best results on the high end of the range.

0.1 M sodium acetate, 0.01 M $ZnCl_2$, 7–10 % w/v PEG 6000, pH 5.0

Crystals formed over night and continued to grow for up to one day and were suitable for soaking of the ligands ATP (10 mM), PRPP (10 mM) and PR-ATP (14 mM) in a 1:2 mixture with 30–60 minutes incubation time. A misture of 20 % w/v PEG 400 and 80 % condition was used as cryoprotectant.

The *Cje*ATP-PRT ACT mutant (4 mg mL $^{-1}$ average drop concentration) formed crystals in JCSG*plus premier*™ HT-96 screen condition E1: 1.0 M sodium citrate tribasic dihydrate, 0.1 M sodium cacodylate, pH 6.5. Crystals grown from this condition were cryoprotected with a mixture of 50 % saturated sucrose and 50 % reservoir and delivered diffraction data out to 4 Å.

The *Cje*ATP-PRT ACT crystallised in a large variety of other screen conditions: JCSG*plus premier*™ HT-96 B9, G10 and PACT *premier*™ HT-96 A4, C7, E6 and G1. Of these only B9 (0.1 M sodium citrate, 20 % w/v PEG 6000, pH 5.0) was able to be up-scaled with reasonable reproducibility, but were of very poor diffraction quality.

6.4 Methods for Chapter 4

6.4.1 Cloning of *Cje*ATP-PRT single site mutants

All single site mutants mentioned in this chapter were created by applying the general SDM protocol described previously (Section 6.1.15). The pDEST™17 vector carrying the *Cje*ATP-PRT gene was used as the template and a temperature range from 45–65 °C was employed during the annealing step of the SDM. All mutagenesis primers are listed below with the mutated base pairs highlighted in bolt script.

Name	**Sequence** (5' - 3')
R54A FWD	CATAGATATTTTACGCGTT**GC**AGATGATGATATCCCTGG
R54A REV	CCAGGGATATCATCATCT**GC**AACGCGTAAAATATCTATG
R216A FWD	GTGGATAAAATTATGTTG**GC**TGTAGCAGGAGTAATGC
R216A REV	GCATTACTCCTGCTACA**GC**CAACATAATTTTATCCAC
R224A FWD	GTAGCAGGAGTAATGCAAGCA**GC**AGAATCTAAATACAT TATGC
R224A REV	GCATAATGTATTTAGATTCT**GC**TGCTTGCATTACTCCT GCTAC
H33A H35S FWD	GAATGTGGTGTTAAAATG**GC**TATC**TC**TGAACAAAGCT TGATCG
H33A H35S REV	CGATCAAGCTTTGTTCA**GA**GATA**GC**CATTTTAACACC ACATTC
H33A H35S L FWD	GAATGTGGTGTTAAAATG**GC**TATC**TC**TGAACAAAGCT TGATCGCTTTTTC
H33A H35S L REV	GAAAAAGCGATCAAGCTTTGTTCA**GA**GATA**GC**CATTT TAACACCACATTC

The resulting reaction mixtures were first screened for the presence of suitable amounts of vector by agarose gel electrophoresis and then transformed into chemically competent *E. coli* TOP10 cells. Two or more clones were picked for each mutant and grown overnight at 37 °C and 180 rpm before plasmid extraction. The purified vectors were sequence analysed to confirm the introduced mutations and transformed into *E. coli* expression strains (BL21* (DE3), BL21* (DE3) pLysS and Chaperone 3) thereafter.

6.4.2 Expression trials

The mutant enzymes were expressed in small test cultures (200 mL volume) for 2 hours at 37 °C and 180 rpm after induction with 0.5 mM IPTG during the logarithmic phase. Cells were harvested and lysed as described (section 6.1.22). His SpinTrap™ single use columns (GE Healthcare) were used according to the manual for fast IMAC binding tests of a small fraction (0.5–1.0 mL) of each clarified supernatant. The solubility and approximate yields from the used expression strains were compared by SDS-PAGE analysis and the eluates of the His SpinTrap™ runs were tested for activity under standard assay conditions (section 6.1.33).

6.4.3 Determination of the kinetic properties of the R216A mutant

Michaelis-Menten constants and turn over rates for *Cje*ATP-PRT R216A were measured under standard assay conditions (section 6.1.33) altering the PRPP concentration (6 μM, 12 μM, 24 μM, 60 μM, 120 μM, 240 μM and 360 μM) and ATP concentration (0.01 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.4 mM, 1 mM, 2 mM and 4 mM) separately. The used enzyme concentration in the cuvette was consistently 6.9 μM.

The dose response curve to histidine was measured under standard assay conditions at 0, 0.2, 2 and 20 mM histidine concentrations.

The dose response curve to AMP was also measured under standard assay conditions with AMP concentrations in the cuvette being 0 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, 2 mM and 5 mM.

6.4.4 ITC

Histidine binding experiments on the *Cje*ATP-PRT R216A mutant were performed as described in section 6.1.39 using an enzyme concentration of 100

μM in the cell and 1.5 mM histidine as titrant.

6.4.5 Crystallisation conditions

All procedures were performed as described in section 6.1.44. The *Cje*ATP-PRT R216A and R54A mutants were crystallised in the following conditions after manual adjustment of the initial *Cje*ATP-PRT conditions. The used protein concentration ranged between 1.5 and 2.5 mg mL $^{-1}$ in the drop comparable to the concentrations used for the CjeATP-PRT wild type.

0.1 M Tris/HCl, 0.2 M KSCN, 15–16 % w/v PEG 4000, 1 mM histidine, pH 7.5

0.1 M Tris/HCl, 0.2 M KSCN, 8 % PEG 1000/ 8 % PEG 8000, 10 mM ATP, pH 7.5

0.1 M Tris/HCl, 0.1 M $MgCl_2$, 13–15 % w/v PEG 4000, 5 mM AMP, pH 7.5

Appendices

Appendix A

Raw kinetic data as proof of *E. coli* pyrophosphatase function

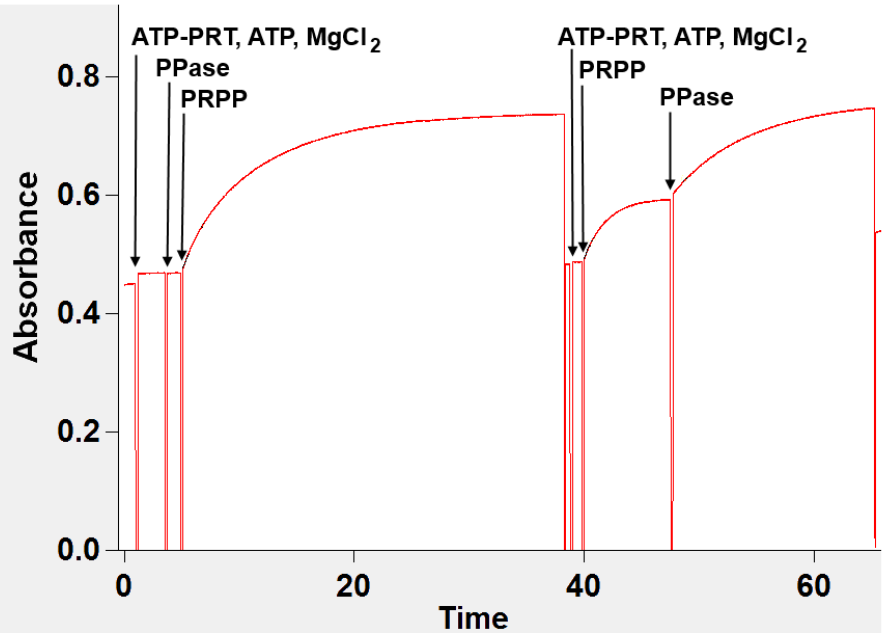


Figure A.1: Raw absorbance data showing the equilibrium character of the *Cje*ATP-PRT reaction. The reaction can be driven to completeness in the biosynthetic direction by the addition of *Eco*PPase.

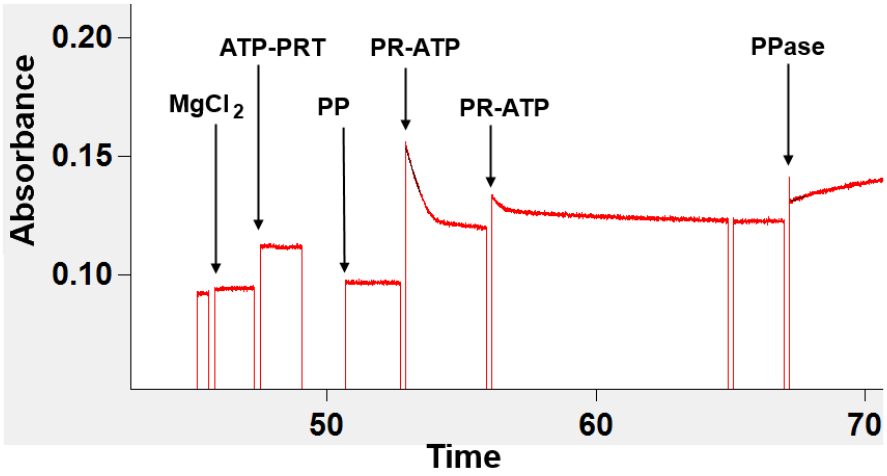


Figure A.2: Raw absorbance data showing the inversion of the *Cje*ATP-PRT reaction by addition of *Eco*PPase.

Appendix B

Michaelis-Menten plots for the determination of K_M values calculated for *Cje*ATP-PRT mutants

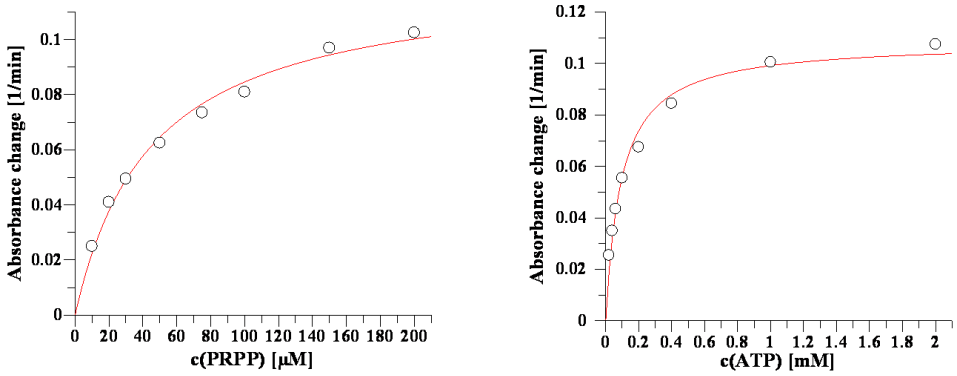


Figure B.1: Michaelis-Menten plots used for determination of the K_M values of PRPP and ATP for *Cje*ATP-PRT Core. K_M and k_{cat} values were determined according to section 6.1.33.

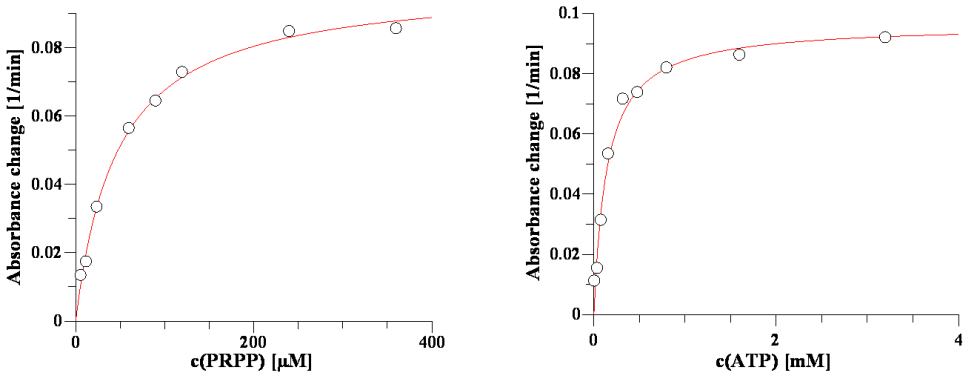


Figure B.2: Michaelis-Menten plots used for determination of the K_M values of PRPP and ATP for *Cje*ATP-PRT R216A. K_M and k_{cat} values were determined according to section 6.1.33.

Appendix C

Michaelis-Menten plots for the determination of K_i values calculated for *Cje*ATP-PRT

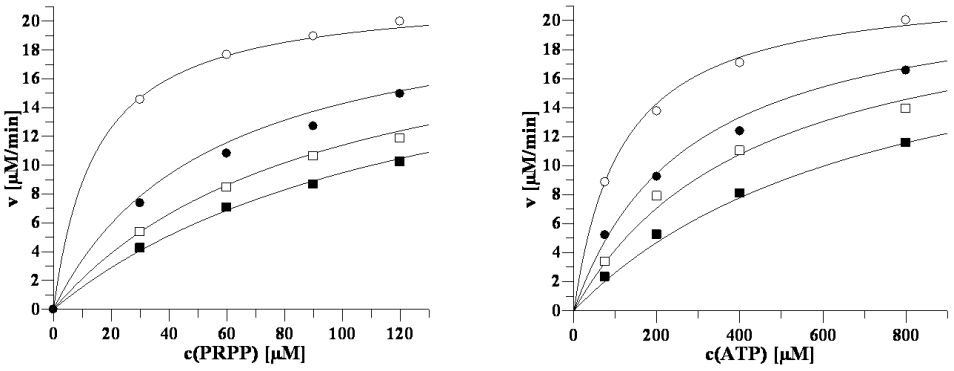


Figure C.1: Michaelis-Menten plots used for determination of the K_i values of AMP against PRPP and ATP for *Cje*ATP-PRT. K_i values were determined by global fitting of the data to the equation for competitive inhibition using GraFit (version 5.0.13, Erithacus Software Limited).

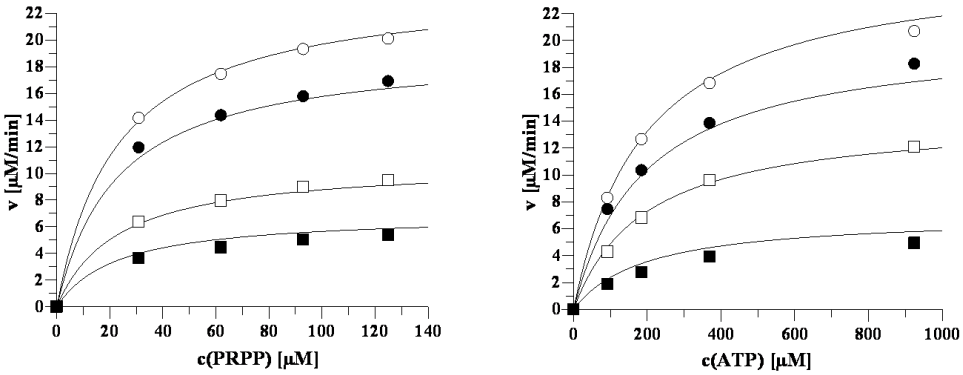


Figure C.2: Michaelis-Menten plots used for determination of the K_i values of histidine against PRPP and ATP for *Cje*ATP-PRT. K_i values were determined by global fitting of the data to the equation for non-competitive inhibition using GraFit (version 5.0.13, Erithacus Software Limited).

Appendix D

Guinier analysis and pair distribution functions for *Cje*ATP-PRT

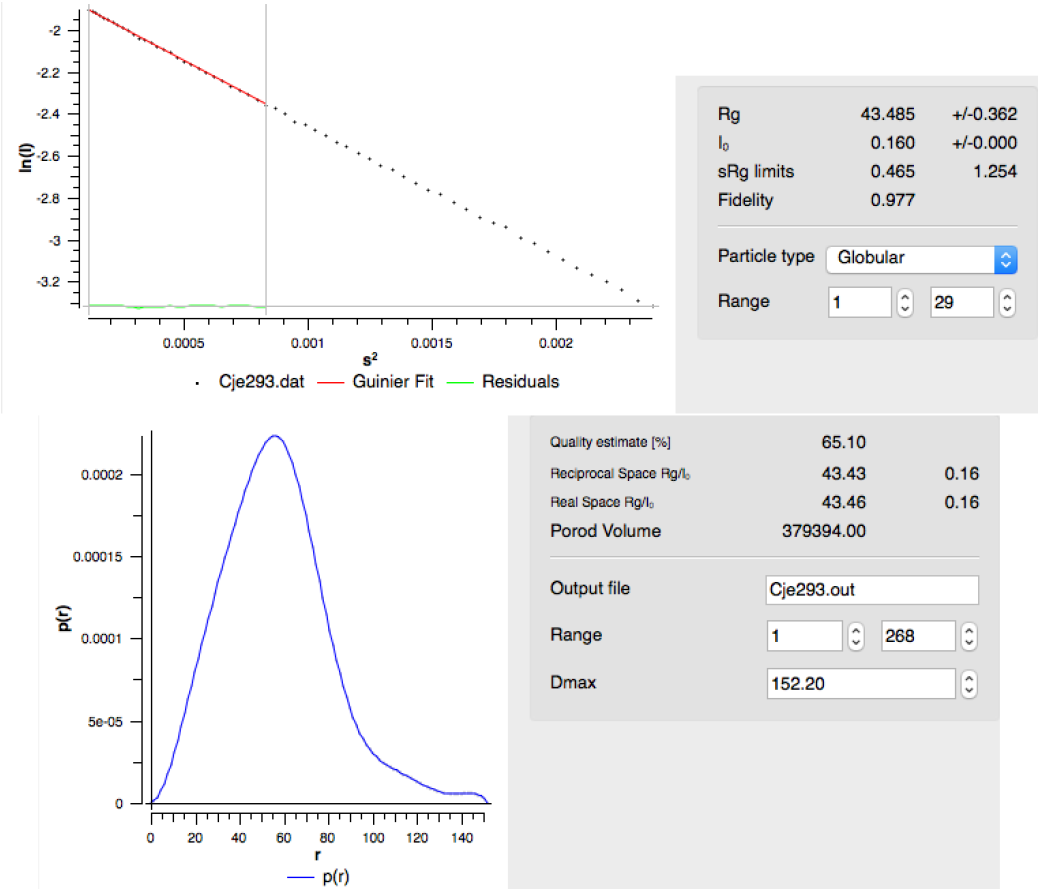


Figure D.1: Guinier analysis and pair distribution functions for *Cje*ATP-PRT wild type. Plots were generated by PRIMUS[107] and GNOM.[95]

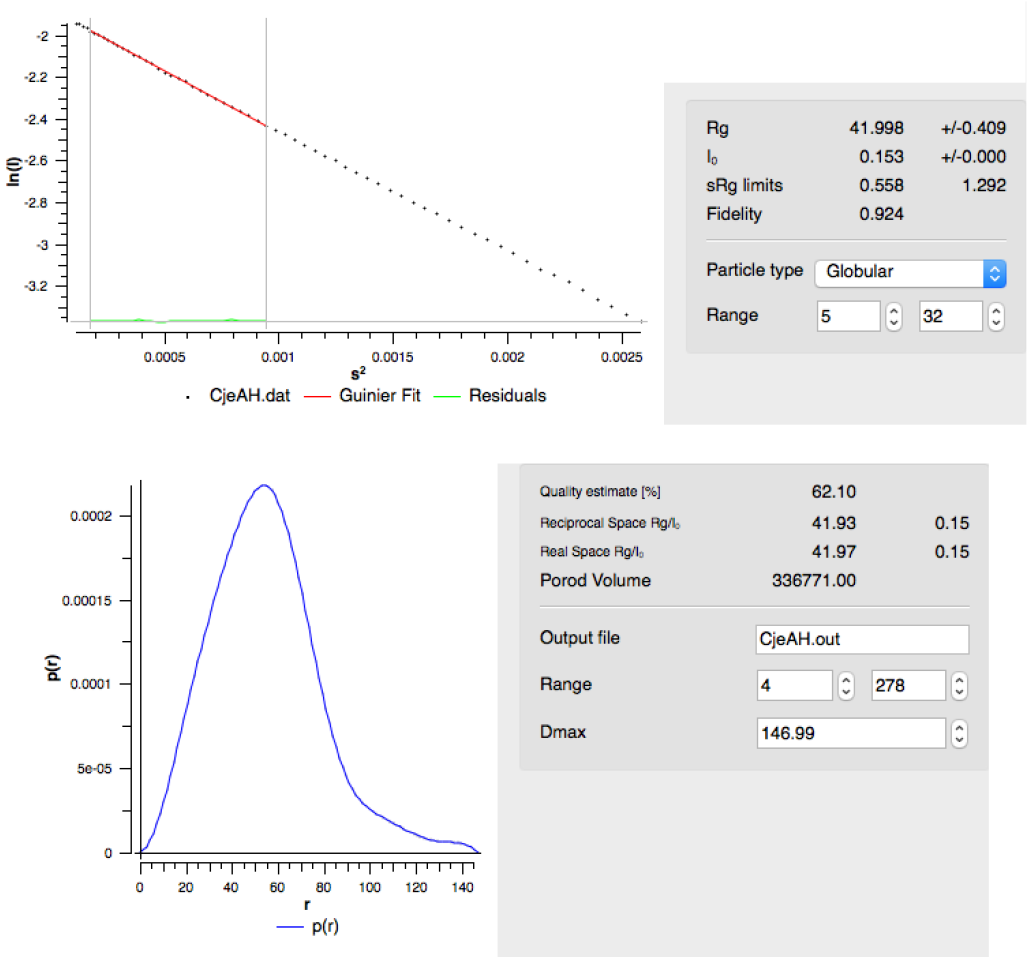


Figure D.2: Guinier analysis and pair distribution functions for *Cje*ATP-PRT wild type in the presence of 1 mM Histidine and 2 mM AMP. Guinier plots were calculated by PRIMUS.[107] Pair wise distribution function P(r) was generated using GNOM.[95]

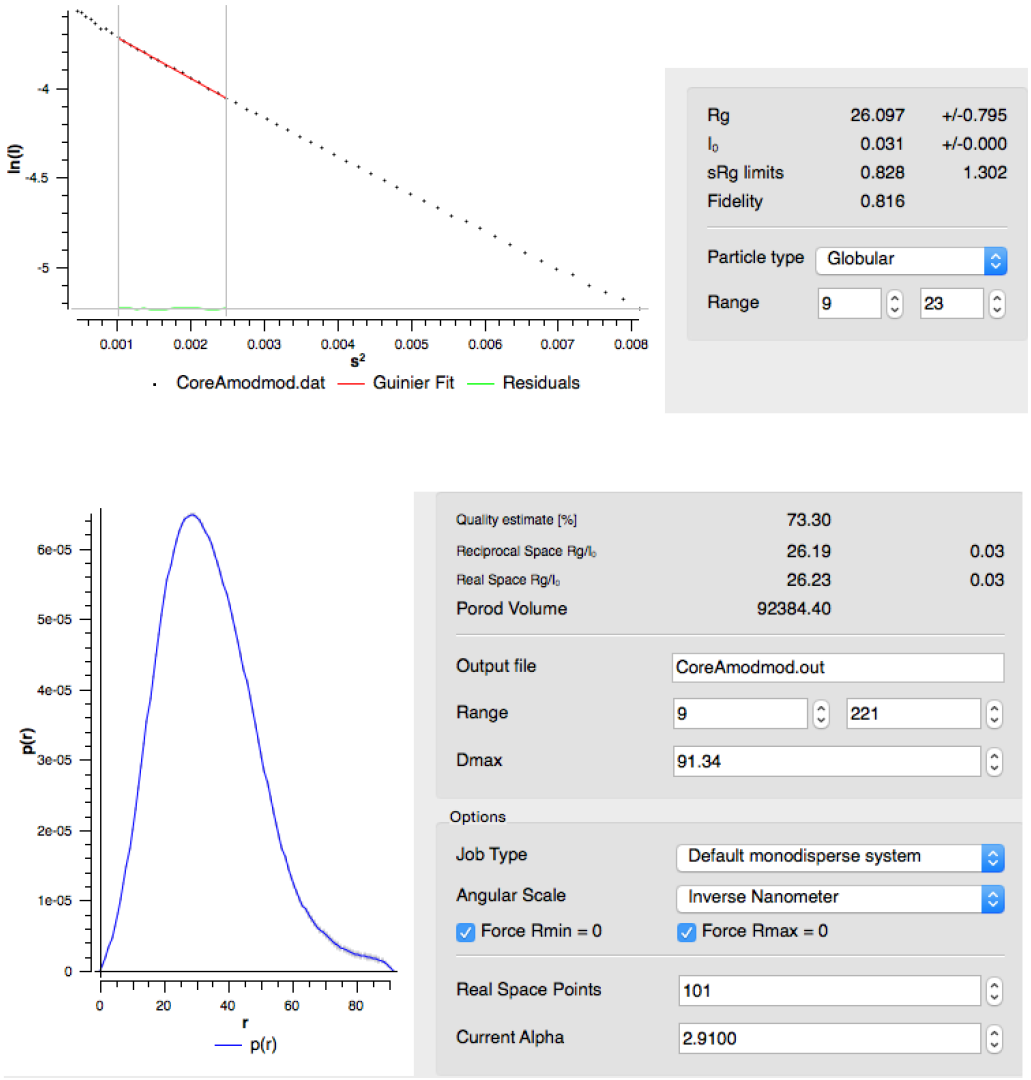


Figure D.3: Guinier analysis and pair distribution functions for *Cje*ATP-PRT Core. Guinier plots were calculated by PRIMUS.[107] Pair wise distribution function P(r) was generated using GNOM.[95]

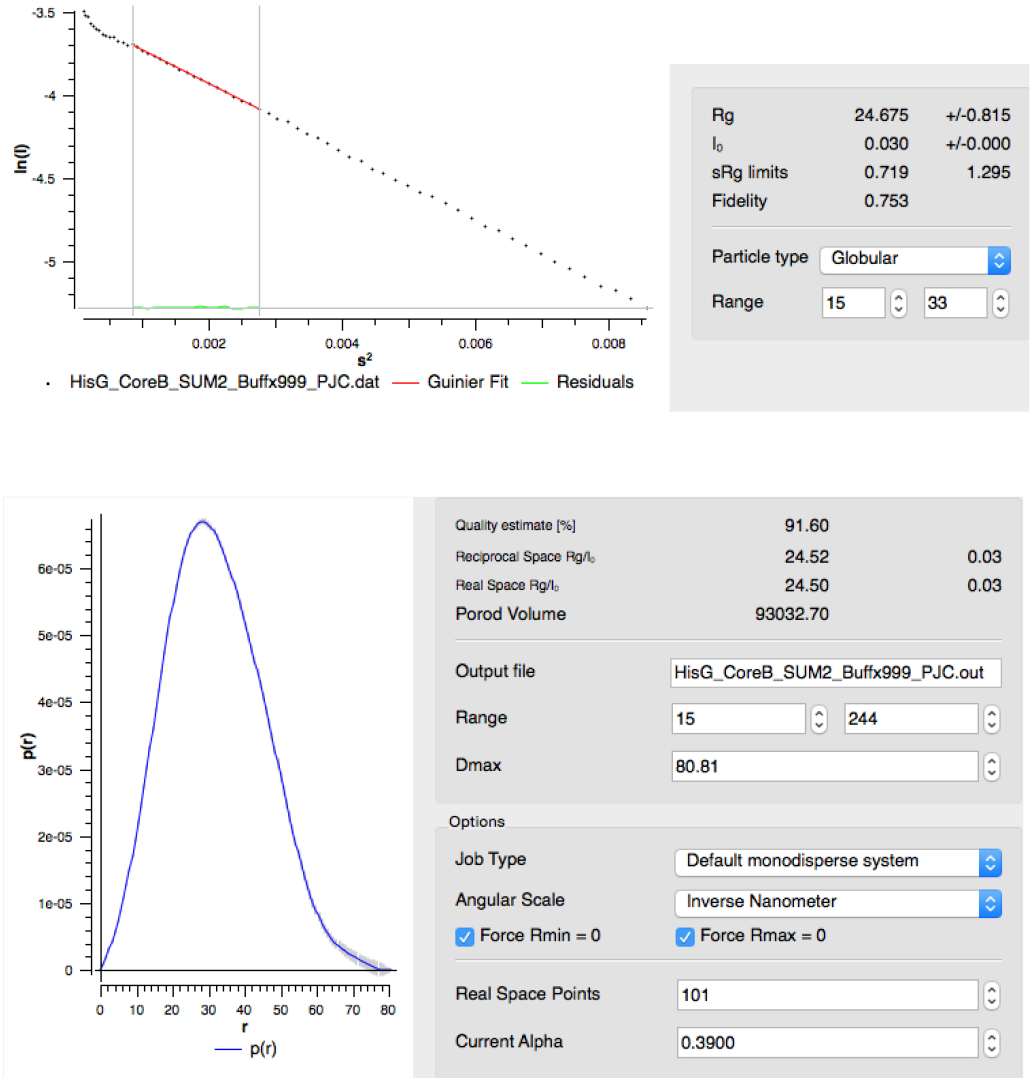


Figure D.4: Guinier analysis and pair distribution functions for *Cje*ATP-PRT Core in the presence of 2 mM AMP. Guinier plots were calculated by PRIMUS.[107] Pair wise distribution function P(r) was generated using GNOM.[95]

References

[1] G. Dodson and A. Wlodawer. Catalytic triads and their relatives. *Trends in Biochemical Sciences*, **9**:347–52, 1998.

[2] B. Sarkar. Metal protein interactions. *Progress in Food and Nutrition Science*, **11**:363–400, 1987.

[3] P. Chakrabarti. Geometry of interaction of metal ions with histidine residues in protein structures. *Protein Engineering Design and Selection*, **4**(1):57–63, 1990.

[4] D. M. Mousdale and J. R. Coggins. Amino acid synthesis. In *Target Sites for Herbicide Action* (ed. R.C. Kirkwood), pp. 29–56. Plenum Press, New York, 1991.

[5] J. Lunardi, J. E. S. Nunes, C. V. Bizarro, L. A. Basso, D. S. Santos, and P. Machado. Targeting the Histidine Pathway in *Mycobacterium tuberculosis*. *Current Topics in Medicinal Chemistry*, **13**:2866–2884, 2013.

[6] F. Haas, M. B. Mitchell, B. N. Ames, and H. K. Mitchell. A series of histidineless mutants of *Neurospora crassa*. *Genetics*, **37**:217–226, 1952.

[7] B.N. Ames, B. Garry, and L. A. Herzenberg. The genetic control of the enzymes of histidine biosynthesis in *Salmonella typhimurium*. *Journal of General Microbiology*, **22**:369–378, 1960.

[8] L. Klungsöyr and D. E. Atkinson. Regulatory properties of phosphoribosyladenosine triphosphate synthetase. Synergism between adenosine monophosphate, phosphoribosyladenosine triphosphate, and histidine. *Biochemistry*, **9**:2021–2027, 1970.

[9] A. Stepansky and T. Leustek. Histidine biosynthesis in plants. *Amino Acids*, **30**:127–142, 2006.

[10] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, **3**:318–356, 1961.

[11] D. Zipser. Polar mutations and operon function. *Nature*, **221**:21–25, 1969.

[12] F. Blasi and C. B. Bruni. Regulation of the histidine operon: translation controlled transcription termination (a mechanism common to several biosynthetic operons). *Current Topics in Cellular Regulation*, **19**:1–45, 1981.

[13] R. G. Martin. The first enzyme in histidine biosynthesis: the nature of feedback inhibition by histidine. *Journal of Biological Chemistry*, **238**:257–268, 1963.

[14] P. Alifano, R. Fani, P. Lio, A. Lazcano, M. Bazzicalupo, M. S. Carlomagno, and C. B. Bruni. Histidine Biosynthetic Pathway and Genes: Structure, Regulation, and Evolution. *Microbiological Reviews*, **60**:44–69, 1996.

[15] T. L. Legerton and C. Yanofsky. Cloning and characterization of the multifunctional his-3 gene of *Neurospora crassa*. *Gene*, **39**:129–140, 1985.

[16] B. N. Ames, R. G. Martin, and B. J. Garry. The First Step of Histidine Biosynthesis. *Journal of Biological Chemistry*, **236**:2019–2026, 1961.

[17] J. E. Kleemann and S. M. Parsons. Reverse Direction Substrate Kinetics and Inhibition Studies on the First Enzyme of Histidine Biosynthesis, Adenosine Triphosphate Phosphoribosyltransferase. *Archives of Biochemistry and Biophysics*, **175**:687–693, 1976.

[18] D. P. Morton and S. M. Parsons. Inhibition of ATP Phosphoribosyltransferase by AMP and ADP in the Absence and Presence of Histidine. *Archives of Biochemistry and Biophysics*, **181**:643–648, 1977.

[19] S. Allen, J.L. Zilles, and D.M. Downs. Metabolic flux in both the purine mononucleotide and histidine biosynthetic pathways can influence synthesis of the hydroxymethyl pyrimidine moiety of thiamine in *Salmonella enterica*. *Journal of Bacteriology*, **184**:6130–6137, 2002.

[20] W. B. Holmes and Appling D. R. Cloning and characterization of methenyltetrahydrofolate synthetase from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, **277**:20205–20213, 2002.

[21] M. Brenner and B. N. Ames. The histidine operon and its regulation. In *Metabolic Regulation,* Vol. 5 of *Metabolic Pathways* (Vogel, H. J., ed.), pp. 349–387. Academic Press Inc, New York, 1971.

[22] S. C. Sinha and J. L. Smith. The PRT protein family. *Current Opinion in Structural Biology*, **11**:733–739, 2001.

[23] J. C. Eads, D. Ozturk, T. B. Wexler, C. Grubmeyer, and J. C. Sacchettini. A new function for a common fold: the crystal structure of quinolinic acid phosphoribosyltransferase. *Structure*, **5**:47–58, 1997.

[24] O. Mayans, A. Ivens, L. J. Nissen, K. Kirschner, and M. Wilmans. Structural analysis of two enzymes catalysing reverse metabolic reactions implies common ancestry. *EMBO Journal*, **21**:3245–3254, 2002.

[25] C. E. Lee, C. Goodfellow, F. Javid-Majd, E. N. Baker, and J. S. Lott. The Crystal Structure of TrpD, a Metabolic Enzyme Essential for Lung Colonization by *Mycobacterium tuberculosis*, in Complex with its Substrate Phosphoribosylpyrophosphate. *Journal of Molecular Biology*, **355**:784–797, 2006.

[26] B. Lohkamp, G. McDermott, S. A. Campbell, J. R. Coggins, and A. J. Lapthorn. The Structure of *Escherichia coli* ATP-phosphoribosyltransferase: Identification of Substrate Binding Sites and Mode of AMP Inhibition. *Journal of Molecular Biology*, **336**:131–144, 2004.

[27] K. S. Champagne, E. Piscitelli, and C. S. Francklyn. Substrate recognition by the hetero-octameric ATP phosphoribosyltransferase from *Lactococcus lactis*. *Biochemistry*, **45**:14933–14943, 2006.

[28] D. P. Morton and S. M. Parsons. Biosynthetic Direction Substrate Kinetics and Product Inhibition Studies on the First Enzyme of Histidine Biosynthesis, Adenosine Triphosphate Phosphoribosyltransferase. *Archives of Biochemistry and Biophysics*, **175**:677–686, 1976.

[29] D. Chelsky and S. M. Parsons. Stereochemical course of the adenosine triphosphate phosphoribosyltransferase reaction in histidine biosynthesis. *Journal of Biological Chemistry*, **250**:5669–5673, 1975.

[30] E. A. T. Ringia and V. L. Schramm. Transition states and inhibitors of the purine nucleoside phosphorylase family. *Current Topics in Medicinal Chemistry*, **5**:1237–1258, 2005.

[31] R. K. Goitein, D. Chelsky, and S. M. Parsons. Primary carbon-14 and α secondary tritium substrate kinetic isotope effects for some phosphoribosyltransferases. *Journal of Biological Chemistry*, **253**:2963–2971, 1978.

[32] J. M. Krahn, J. H. Kim, M. R. Burns, R. J. Parry, H. Zalkin, and J. L. Smith. Coupled Formation of an Amidotransferase Interdomain Ammonia Channel and a Phosphoribosyl-transferase Active Site. *Biochemistry*, **36**:11061–11068, 1997.

[33] J.L. Smith. Forming and inhibiting PRT active sites. *Nature Structural and Molecular Biology*, **6**:502–504, 1999.

[34] T. V. Cookson, A. Castell, E. M. Bulloch, G. L. Evans, F. L. Short, E. N. Baker, J. S. Lott, and E. J. Parker. Alternative substrates reveal catalytic cycle and key binding events in the reaction catalysed by anthranilate phosphoribosyltransferase from *Mycobacterium tuberculosis*. *The Biochemical Journal*, **461**:87–98, 2014.

[35] Y. Cho, V. Sharma, and J. C. Sacchettini. Crystal Structure of ATP Phosphoribosyltransferase from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, **278**(10):8333–8339, 2003.

[36] K. S. Champagne, M. Sissler, Y. Larrabee, S. Doublié, and C. S. Francklyn. Activation of the Hetero-octameric ATP Phosphoribosyl Transferase through Subunit Interface Rearrangement by a tRNA Synthetase Paralog. *Journal of Biological Chemistry*, **280**(40):34096–34104, 2005.

[37] M. C. Vega, P. Zou, F. J. Fernandez, G. E. Murphy, R. Sterner, A. Popov, and M. Wilmanns. Regulation of the hetero-octameric ATP phosphoribosyl transferase complex from *Thermotoga maritima* by a tRNA synthetase-like subunit. *Molecular Microbiology*, **55**(3):675–686, 2005.

[38] D. Schuller, G. A. Grant, and L. Banaszak. The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. *Nature Structural and Molecular Biology*, **2**:69–76, 1995.

[39] L. Aravind and E. V. Koonin. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *Journal of Molecular Biology*, **287**:1023–1040, 1999.

[40] G.A. Grant. The ACT Domain: A Small Molecule Binding Domain and Its Role as a Common Regulatory Element. *Journal of Biological Chemistry*, **281**:33825–33829, 2002.

[41] E. J. M. Lang, P. J. Cross, G. Mittelstädt, G. B. Jameson, and E. J. Parker. Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Current Opinion in Structural Biology*, **29**:102–111, 2014.

[42] R. M. Bell, S. M. Parsons, S. A. Dubravac, A. G. Redfield, and D. E. Jr. Koshland. Characterization of Slowly Interconvertible States of Phosphoribosyladenosine Triphosphate Synthetase Dependent on Temperature, Substrates, and Histidine. *Journal of Biological Chemistry*, **249**:4110–4118, 1974.

[43] A. R. Tebar, V. M. Fernandez, R. Martin Del Rio, and A. O. Ballesteros. Studies on the quaternary structure of the first enzyme for histidine biosynthesis. *Experientia*, **29**:1477–1479, 1973.

[44] B. Lohkamp, J. R. Coggins, and A. J. Lapthorn. Purification, crystallization and preliminary X-ray crystallographic analysis of ATP-phosphoribosyltransferase from *Escherichia coli*. *Acta Crystallographica, Section D: Biological Crystallography*, D**56**:1488–1491, 2000.

[45] M. Sissler, C. Delorme, J. Bond, S. D. Ehrlich, P. Renault, and C. S. Francklyn. An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **96**:8985–8990, 1999.

[46] N. Yao, S. Trakhanov, and F. A. Quiocho. Refined 1.89-A structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry*, **33**(16):4769–4779, 1994.

[47] B. C. H. Chu, T. DeWolf, and H. J. Vogel. Role of the Two Structural Domains from the Periplasmic *Escherichia coli* Histidine-binding Protein HisJ. *Journal of Biological Chemistry*, **288**:31409–31422, 2013.

[48] L. F. Huergo, G. Chandra, and M. Merrick. P_{II} signal transduction proteins: nitrogen regulation and beyond. *FEMS Microbiology Reviews*, **37**:251–283, 2013.

[49] V.R. Chellamuthu, V. Alva, and K. Forchhammer. From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta*, **237**:451–462, 2012.

[50] M. Aklujkar. Two ATP phosphoribosyltransferase isozymes of *Geobacter sulfurreducens* contribute to growth in the presence or absence of histidine and under nitrogen fixation conditions. *Canadian Journal of Microbiology*, **57**:547–558, 2011.

[51] J. P. Bond and C. Francklyn. Proteobacterial Histidine-Biosynthetic Pathways Are Paraphyletic. *Journal of Molecular Evolution*, **50**:339–347, 2000.

[52] C. Francklyn. tRNA synthetase paralogs: Evolutionary links in the transition from tRNA-dependent amino acid biosynthesis to *de novo* biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(17):9650–9652, 2003.

[53] J.M. Berg, J.L. Tymoczko, and L. Stryer. *Biochemistry*. W. H. Freeman, 2010.

[54] Y. Cheng, Y. Zhou, L. Yang, C. Zhang, Q. Xu, X. Xie, and N. Chen. Modification of histidine biosynthesis pathway genes and the impact on production of L-histidine in *Corynebacterium glutamicum*. *Biotechnology Letters*, **35**:735–741, 2013.

[55] C. Yanofsky. Attenuation in the control of expression of bacterial operons. *Nature*, **289**:751–758, 1981.

[56] R. Kolter and C. Yanofsky. Attenuation in Amino Acid Biosynthetic Operons. *Annual Review in Genetics*, **16**:113–134, 1982.

[57] W. W. Cleland. The kinetics of enzyme-catalyzed reactions with two or more substrates or products. II. Inhibition: nomenclature and theory. *Biochimica et Biophysica Acta*, **67**:173–187, 1963.

[58] J. P. Changeux. 50th anniversary of the word "allosteric". *Protein Science*, **20**:1119–1124, 2011.

[59] G. Suel, S. Lockless, M. Wall, , and R. Ranganathan. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Structural and Molecular Biology*, **10**:59–69, 2003.

[60] C. Tsai, A. del Sol, , and R. Nussinov. Allostery: Absence of a change in shape does not imply that allostery is not at play. *Journal of Molecular Biology*, **378**:1–11, 2008.

[61] J. O. Wrabl, J. Gu, T. Liu, T. P. Schrank, S. T. Whitten, and V. J. Hilser. The role of protein conformational fluctuations in allostery, function, and evolution. *Biophysical Chemistry*, **159**:129–141, 2011.

[62] C. Bohr. Die Sauerstoffaufnahme des genuinen Blutfarbstoffes und des aus dem Blute darstellten Hämoglobins. *Centralblatt für Physiologie*, **23**, 1904.

[63] A. Hill. The possible effect of the aggregation of the molecules of haemoglobin on its dissociation curve. *Journal of Physiology*, **40**:4–7, 1910.

[64] H. E. Umbarger. Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science*, **123**:848, 1956.

[65] H. E. Umbarger. Feedback control by endproduct inhibition. *Cold Spr. Harb. Symp. Quant. Biol.*, **26**:301–312, 1961.

[66] J. Monod, J. P. Changeux, and F. Jacob. Allosteric proteins and cellular control systems. *Journal of Molecular Biology*, **6**:306–329, 1963.

[67] J. Monod, J. Wyman, and J. P. Changeux. On nature of allosteric transitions — A plausible model. *Journal of Molecular Biology*, **12**:88–118, 1965.

[68] D. E. Jr Koshland, G. Némethy, and D. Filmer. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry*, **5**:365–385, 1966.

[69] J. P. Changeux and M. M. Rubin. Allosteric interactions in aspartate transcarbamylase. III. Interpretations of experimental data in terms of the model of Monod, Wyman, and Changeux. *Biochemistry*, **7**:553–561, 1968.

[70] E. K. Jaffe. Morpheeins — a new structural paradigm for allosteric regulation. *TRENDS in Biochemical Sciences*, **30**:490–497, 2005.

[71] D. E. Koshland. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **44**:98–104, 1958.

[72] B. Volkman, D. Lipson, D. Wemmer, and D. Kern. Two-state allosteric behavior in a single-domain signaling protein. *Science*, **291**:2429–2433, 2001.

[73] D.W. Gohara and E. Di Cera. Allostery in trypsin-like proteases suggests new therapeutic strategies. *Trends in Biotechnology*, **29**:577–585, 2011.

[74] H. Kryvi and L. Klungsöyr. Kinetic properties of phosphoribosyladenosine triphosphate synthetase. Inhibition by aggregation at high enzyme concentrations. *Biochemica et Biophysica Acta*, **235**(3):429–434, 1971.

[75] Y. Cho, T. R. Ioerger, and J. C. Sacchettini. Discovery of novel nitrobenzothiazole inhibitors for *Mycobacterium tuberculosis* ATP phosphoribosyl transferase (HisG) through virtual screening. *Journal of Medical Chemistry*, **51**(19):5984–5992, 2012.

[76] L. Klungsöyr and H. Kryvi. Sedimentation behaviour of phosphoribosyladenosine triphosphate synthetase. Effects of substrates and modifiers. *Biochemica et Biophysica Acta*, **227**(2):327–336, 1971.

[77] Y. Zhang, X. Shang, A. Deng, X. Chai, S. Lai, G. Zhang, and T. Wen. Genetic and biochemical characterization of *Corynebacterium glutamicum* ATP phosphoribosyltransferase and its three mutants resistant to feedback inhibition by histidine. *Biochimie*, **94**:829–838, 2012.

[78] P. Vandamme, F. E. Dewhirst, B. J. Paster, and S. L. W. On. *Bergey's Manual of Systematic Bacteriology: Volume Two: The Proteobacteria (Part C)* (2nd ed.), pp. 1147–1160. Springer Science and Business Media, 2006.

[79] J. I. Keller and W. G. Shriver. Prevalence of three Campylobacter Species, *C. jejuni*, *C. coli*, and *C. lari*, using Multilocus Sequence Typing in Wild Birds of the Mid-Atlantic Region, USA. *Journal of Wildlife Diseases*, **50**(1):31–41, 2014.

[80] B. M. Allos. *Campylobacter jejuni* Infections: Update on Emerging Issues and Trends. *Clinical Infectious Diseases*, **32**(8):1201–1206, 2001.

[81] R. M. Humphries and A. J. Linscott. Laboratory Diagnosis of Bacterial Gastroenteritis. *Clinical Microbiology Reviews*, **28**(1):3–31, 2015.

[82] J. Parkhill, B. W. Wren, K. Mungall, J. M. Ketley, C. Churcher, D. Basham, T. Chillingworth, R. M. Davies, T. Feltwell, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Moule, M. J. Pallen, C. W. Penn, M. A. Quail, M-A. Rajandream, K. M. Rutherford, A. H. M. van Vliet, S. Whitehead, and B. G. Barell. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature*, **403**:665–668, 2000.

[83] M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Research*, **42**:D199–D205, 2014.

[84] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, Lopez R., H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, **7**:539, 2011.

[85] A. Rambaut. FigTree, http://tree.bio.ed.ac.uk/software/figtree/.

[86] T. Cavalier-Smith. A revised six-kingdom system of life. *Biological Reviews*, **73**:203–266, 1998.

[87] D. Ohta, K. Fujimori, M. Mizutani, Y. Nakayama, R. Kunpaisal-Hashimo, S. Münzer, and A. Kozaki. Molecular Cloning and Characterization of ATP-Phosphoribosyl Transferase from Arabidopsis, a Key Enzyme in the Histidine Biosynthetic Pathway. *Plant Physiology*, **122**:907–914, 2000.

[88] X. Robert and P. Gouet. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, **42**(W1):W320–W324, 2014.

[89] S. Pedreño, J. P. Pisco, G. Larrouy-Maumus, G. Kelly, and L. P. S. de Carvalho. Mechanism of feedback allosteric inhibition of ATP phosphoribosyltransferase. *Biochemistry*, **51**:8027–8038, 2012.

[90] P. Schuck. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal*, **78**:1606–1619, 2000.

[91] S. T. Rao and M. G. Rossmann. Comparison of super-secondary structures in proteins. *Journal of Molecular Biology*, **76**(2):241–256, 1973.

[92] A. Stivala, M. Wybrow, A. Wirth, J. Whisstock, and P. Stuckey. Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics*, **27**(23):3315–3316, 2011.

[93] E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, **372**:774–797, 2007.

[94] R. A. Laskowski and M. B. Swindells. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling*, **51**:2778–2786, 2011.

[95] D. I. Svergun. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, **25**:495–503, 1992.

[96] D. I. Svergun, C. Barberato, and Koch M. H. J. CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, **28**:768–773, 1995.

[97] E. K. Livingstone. *Allosteric Regulation of the First Enzyme in Histidine Biosynthesis*. Master's thesis, University of Canterbury, NZ, 2015.

[98] R. K. Kulis-Horn, M. Persicke, and J. Kalinowski. Corynebacterium glutamicum ATP-phosphoribosyl transferases suitable for L-histidine production — Strategies for the elimination of feedback inhibition. *Journal of Biotechnology*, **206**:26–37, 2015.

[99] Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.7.4.0. 2014.

[100] B. M. Bulheller and J. D. Hirst. DichroCalc-circular and linear dichroism online. *Bioinformatics*, **25**(4):539–540, 2009.

[101] J. Badger, J. M. Sauder, J. M. Adams, S. Antonysamy, K. Bain, M. G. Bergseid, S. G. Buchanan, M. D. Buchanan, Y. Batiyenko, J. A. Christopher, S. Emtage, A. Eroshkina, I. Feil, E. B. Furlong, K. S. Gajiwala, X. Gao, D. He, J. Hendle, A. Huber, K. Hoda, P. Kearins, C. Kissinger, B. Laubert, H. A. Lewis, J. Lin, K. Loomis, D. Lorimer, G. Louie, M. Maletic, C. D. Marsh, I. Miller, J. Molinari, H. J. Muller-Dieckmann, J. M. Newman, B. W. Noland, B. Pagarigan, F. Park, T. S. Peat, K. W. Post, S. Radojicic, A. Ramos, R. Romero, M. E. Rutter, W. E. Sanderson, K. D. Schwinn, J. Tresser, J. Winhoven, T. A. Wright, L. Wu, J. Xu, and T. J. Harris. Structural analysis of a set of proteins resulting from a bacterial genomics project. *Proteins*, **60**(4):787–796, 2005.

[102] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**(21):2947–2948, 2007.

[103] J. L. Hartley, G. F. Temple, and M. A. Brasch. DNA Cloning Using *in vitro* Site-Specific Recombination. *Genome Research*, **10**:1788–1795, 2000.

[104] U. B. Ericsson, B. M. Hallberg, G. T. DeTitta, N. Dekker, and P. Nordlund. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical Biochemistry*, **357**:289–298, 2006.

[105] P. Schuck, M. S. Perugini, N. R. Gonzales, G. J. Howlett, and D. Schubert. Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophysical Journal*, **82**:1096–1111, 2002.

[106] T. M. Laue, B. D. Shah, T. M. Ridgeway, and S. L. Pelletier. Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (ed. S.E. Harding et al.), pp. 90–125. Royal Society of Chemistry, Cambridge, United Kingdom, 1992.

[107] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, **36**:1277–1282, 2003.

[108] T. M. McPhillips, S. E. McPhillips, H.-J. Chiu, A. E. Cohen, A. M. Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M.

Soltis, and P. Kuhn. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of Synchrotron Radiation*, **9**:401–406, 2002.

[109] W. Kabsch. XDS Program Package, http://xds.mpimf-heidelberg.mpg.de/.

[110] French G. S. and Wilson K. S. On the treatment of negative intensity observations. *Acta Crystallographica, Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, A**34**:517, 1978.

[111] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson. Overview of the CCP4 suite and current developments. *Acta Crystallographica, Section D: Biological Crystallography*, D**67**:235–242, 2011.

[112] A. Vagin and A. Teplyakov. *MOLREP*: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography*, 30:1022–1025, 1997.

[113] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. *Phaser* crystallographic software. *Journal of Applied Crystallography*, **40**:658–674, 2007.

[114] G. Murshudov, A. Vagin, and E. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica, Section D: Biological Crystallography*, D**53**:240–255, 1997.

[115] P. Emsley, B. Lohkamp, W. Scott, and K. Cowtan. Features and Development of Coot. *Acta Crystallographica, Section D: Biological Crystallography*, D**66**:486–501, 2010.

[116] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, and D. C. Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica, Section D: Biological Crystallography*, D**66**:12–21, 2010.